

AD-A150 031

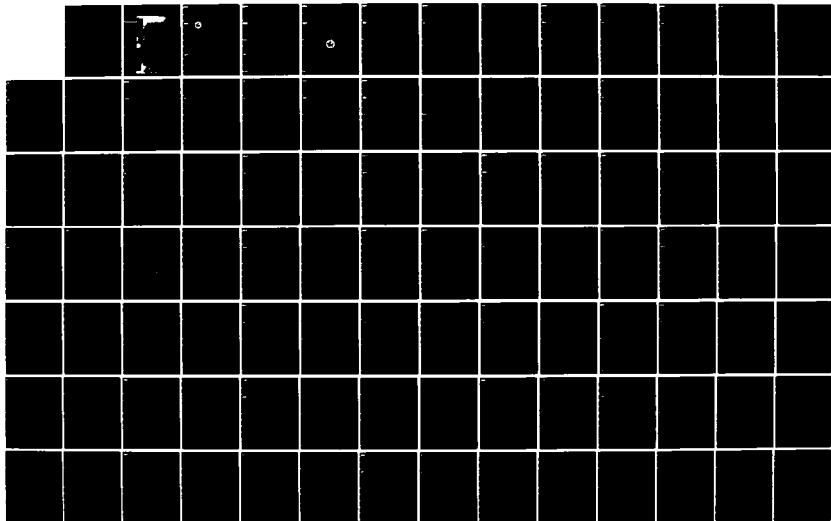
ANNALS OF THE NEW YORK ACADEMY OF SCIENCES VOLUME 423
TIMING AND TIME PER. (U) NEW YORK ACADEMY OF SCIENCES
NY J GIBBON ET AL. 13 MAY 84 N00014-84-G-0130

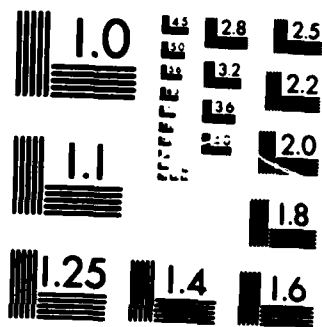
1/3

UNCLASSIFIED

F/G 5/10

NL





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

AD-A150 031

3

TIMING AND TIME PERCEPTION

EDITORS: JOHN GIBSON
LORRAINE ALLAN

DTIC FILE COPY

DTIC
SELECTE

SEP 25 1984

E

This document has been approved
for public release and sale; its
distribution is unlimited.

ANNALS OF THE NEW YORK
ACADEMY OF SCIENCES VOLUME

84 09 06 019

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Volume 423



EDITORIAL STAFF

Executive Editor

BILL BOLAND

Managing Editor

JOYCE HITCHCOCK

Associate Editor

JUSTINE CULLINAN

The New York Academy of Sciences

2 East 63rd Street

New York, New York 10021

THE NEW YORK ACADEMY OF SCIENCES

(Founded in 1817)

BOARD OF GOVERNORS, 1984

CRAIG D. BURRELL, *President*

KURT SALZINGER, *President-Elect*

Honorary Life Governors

SERGE A. KORFF
I. B. LASKOWITZ

H. CHRISTINE REILLY
IRVING J. SELIKOFF

Vice Presidents

WILLIAM S. CAIN
FLORENCE L. DENMARK
PETER M. LEVY
(on leave in 1984)

HARRY LUSTIG
NORBERT J. ROBERTS
FLEUR L. STRAND

ALAN J. PATRICOFF, *Secretary-Treasurer*

Elected Governors-At-Large

1982-84

1984-86

WILLIAM T. GOLDEN
DENNIS KELLY

HERBERT SHEPPARD
DONALD B. STRAUS

MURIEL FEIGELSON

Past Presidents (Governors)

WALTER N. SCOTT

HEINZ R. PAGELS, *Executive Director*

JACQUELINE MESSITE

MORRIS H. SHAMOS

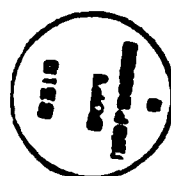
TIMING AND TIME PERCEPTION

Volume 423

TIMING AND TIME PERCEPTION

Edited by John Gibbon and Lorraine Allan

1000-14-84-G-0130



Accession For

NTIS GRA&I ☒

DTIC TAB ☐

Unannounced ☐

Justification

DISTRIBUTION STATEMENT

UNCLASSIFIED

DATE 08-01-97 BY [illegible]

1A-1

*The New York Academy of Sciences
New York, New York
1984*

REPRODUCED AT GOVERNMENT EXPENSE

Copyright © 1984 by The New York Academy of Sciences. All rights reserved. Under the provisions of the United States Copyright Act of 1976, individual readers of the Annals are permitted to make fair use of the material in them for teaching or research. Permission is granted to quote from the Annals provided that the customary acknowledgment is made of the source. Material in the Annals may be republished only by permission of The Academy. Address inquiries to the Executive Editor at The New York Academy of Sciences.

Copying fees: For each copy of an article made beyond the free copying permission under Section 107 or 108 of the 1976 Copyright Act, a fee should be paid through the Copyright Clearance Center, 21 Congress Street, Salem, Mass. 01970. For articles of more than 3 pages, the copying fee is \$1.75.

Library of Congress Cataloging in Publication Data

Main entry under title:

Timing and time perception.

(Annals of the New York Academy of Sciences ; v. 423)

Bibliography: p.

Includes index.

I. Time perception—Congresses. I. Gibbon, John, 1934– . II. Allan, Lorraine, 1940– . III. Series. Q11.N5 vol. 423 [QP445] 500 s [153.7'53] 84-11443

SP

Printed in the United States of America

ISBN 0-89766-240-7 (cloth)

ISBN 0-89766-241-5 (paper)

ISSN 0077-8923

REPRODUCED AT GOVERNMENT EXPENSE

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Volume 423

May 11, 1984

TIMING AND TIME PERCEPTION*

Editors and Conference Co-Chairmen

JOHN GIBBON AND LORRAINE ALLAN

CONTENTS

Part I. Time Perception

Introduction. <i>By</i> JOHN GIBBON AND LORRAINE ALLAN	1
Quantal and Deterministic Timing in Human Duration Discrimination. <i>By</i> ALFRED B. KRISTOFFERSON	3
Ultrastable Stimulus-Response Latencies: Towards a Model of Response-Stimulus Synchronization. <i>By</i> GORDON W. HOPKINS	16
The Perception of Temporal Events. <i>By</i> D. ALAN STUBBS, L. R. DREYFUS, AND J. G. FETTERMAN	30
Subjective Duration in Rats: The Psychophysical Function. <i>By</i> HANNES EISLER	43
Scalar Timing in Memory. <i>By</i> JOHN GIBBON, RUSSELL M. CHURCH, AND WARREN H. MECK	52
Time Perception: Discussion Paper. <i>By</i> R. DUNCAN LUCE	78
The Representation of Mental Activities in Critical Path Networks. <i>By</i> RICHARD J. SCHWEICKERT	82
Timing Perturbations with Complex Auditory Stimuli. <i>By</i> DONALD G. JAMIESON, ELZBIETA SLAWINSKA, MARGARET F. CHEESMAN, AND BLAS ESPINOZA-VARAS	96
Temporal Order and Duration: Their Discrimination and Retention by Pigeons. <i>By</i> EDWARD A. WASSERMAN, ROBERT E. DELONG, AND MARK B. LAREW ..	103
Contingent Aftereffects in Duration Judgments. <i>By</i> LORRAINE G. ALLAN	116
Contingent Aftereffects and Situationally Coded Criteria: Discussion Paper. <i>By</i> MICHEL TREISMAN	131

Part II. Timing of Motor Programs and Temporal Patterns

Introduction. <i>By</i> ALAN M. WING	143
Motor Programming and Temporal Patterns in Handwriting. <i>By</i> GEORGE E. STELMACH, PATRICIA A. MULLINS, AND HANS-LEO TEULINGS	144
The Patterning of Time and Its Effects on Perceiving. <i>By</i> MARI RIESS JONES	158

✓ *This volume is the result of a conference on Timing and Time Perception held on May 10-13, 1983 by The New York Academy of Sciences.

On Controlling Force and Time in Rhythmic Movement Sequences: The Effect of Stress Location. <i>By</i> ANDRAS SEMJEN, ADELA GARCIA-COLERA, AND JEAN REQUIN	168
Motor Disorder and the Timing of Repetitive Movements. <i>By</i> ALAN M. WING, STEVEN KEELE, AND DAVID I. MARGOLIN	183
Timing of Motor Programs and Temporal Patterns: Discussion Paper. <i>By</i> EUGENE GALANTER	193

Part III. Timing in Animal Learning and Behavior

Introduction. <i>By</i> HERBERT S. TERRACE	199
Motivational and Response Factors in Temporal Differentiation. <i>By</i> JOHN R. PLATT	200
Relative Time in Trace Conditioning. <i>By</i> PETER BALSAM	211
The Function of Time Discrimination and Classical Conditioning. <i>By</i> SETH ROBERTS AND MARK D. HOLDER	228
Time and Contingency in Classical Conditioning. <i>By</i> H. M. JENKINS	242
Timing Competence and Timing Performance: A Cross-Species Approach. <i>By</i> MARC RICHELLE AND HELGA LEJEUNE	254
The Integration of Reinforcements over Time. <i>By</i> S. E. G. LEA AND S. M. DOW	269
Time Horizons of Foraging Animals. <i>By</i> JOHN R. KREBS AND ALEJANDRO KACELNIK	278
Timing in Animal Learning and Behavior: Discussion Paper. <i>By</i> EDMUND FANTINO	292

Part IV. Time in Cognitive Processing and Memory

Introduction. <i>By</i> H. L. ROITBLAT	297
Attentional Effort and Cognitive Strategies in the Processing of Temporal Information. <i>By</i> JOHN A. MICHON AND JANET L. JACKSON	298
Time and Memory. <i>By</i> J. E. R. STADDON	322
Temporal Variables in Delayed Matching to Sample. <i>By</i> WILLIAM A. ROBERTS AND PHILIPP J. KRAEMER	335
Timing, Learning, and Forgetting. <i>By</i> CHARLES P. SHIMP	346
A Model for Temporal Generalization and Discrimination. <i>By</i> ERIC G. HEINEMANN	361
Time's Role for Information, Processing, and Normalization. <i>By</i> DOMINIC W. MASSARO	372
Time in Cognitive Processing and Memory: Discussion Paper. <i>By</i> DOUGLAS L. MEDIN	385

Part V. Rhythmic Patterns and Music

Introduction. <i>By</i> JOHN R. PLATT	389
Timing of Two-Handed Rhythmic Performance. <i>By</i> DIRK VORBERG AND ROLF HAMBUCH	390

Processing of Rhythmic Sound Structures by Birds. <i>By</i> STEWART H. HULSE, JOHN HUMPAL, AND JEFFREY CYNX	407
Timing in Musical Performance. <i>By</i> HENRY SHAFFER	420
Perception, Production, and Imitation of Time Ratios by Skilled Musicians. <i>By</i> SAUL STERNBERG AND RONALD L. KNOLL	429
Distinguished Address: Circadian Timing. <i>By</i> JÜRGEN ASCHOFF	442

Part VI. The Internal Clock

Introduction. <i>By</i> RUSSELL M. CHURCH	469
Daily Meal Anticipation: Interaction of Circadian and Interval Timing. <i>By</i> MICHAEL Terman, JOHN GIBBON, STEPHEN FAIRHURST, AND AMY WARING	470
Reproductive Mechanisms: Interaction of Circadian and Interval Timing. <i>By</i> RAE SILVER AND ERIC L. BITTMAN	488
Incentive Theory III: Adaptive Clocks. <i>By</i> PETER R. KILLEEN	515
Attentional Bias between Modalities: Effect on the Internal Clock, Memory, and Decision Stages Used in Animal Time Discrimination. <i>By</i> WARREN H. MECK	528
Temporal Rhythms and Cerebral Rhythms. <i>By</i> MICHEL TREISMAN	542
Properties of the Internal Clock. <i>By</i> RUSSELL M. CHURCH	566

Poster Papers

Retrospective Duration Judgments of a Hypnotic Time Interval. <i>By</i> SCOTT W. BROWN	583
Scalar Timing and the Spatial Organization of Behavior between Reward Presentations. <i>By</i> F. R. CAMPAGNONI AND P. S. COHEN	585
Foraging in the Laboratory: Effects of Session Length. <i>By</i> SUSAN M. DOW	588
Time-Order Discrimination of Sequences of Four Events. <i>By</i> STEPHEN R. GRICE	592
An Endogenous Metric for the Control of Perception of Brief Temporal Intervals. <i>By</i> P. A. HANCOCK	594
The Role of Temporal Factors in Multiple Schedules. <i>By</i> ALISON D. HASSIN-HERMAN	597
Discrimination of Temporal Components of Acoustic Patterns by Birds. <i>By</i> JOHN HUMPAL AND JEFFREY A. CYNX	600
The Processing of Temporal Information. <i>By</i> J. L. JACKSON, J. A. MICHON, AND A. VERMEEREN	603
Effect of Photoperiod on Timing of Incubation in Ring Doves. <i>By</i> ROBERT M. KAHN	605
Duration Perception and Auditory Masking. <i>By</i> HOWARD J. KALLMAN AND MARYELLEN D. MORRIS	608
Familiar Melodies Seem Shorter, Not Longer, When Played Backwards: Data and Theory. <i>By</i> KATHLEEN KOWAL	610
Tactile Temporal Acuties. <i>By</i> EUGENE C. LECHOLT	612

Lick Rate and the Circadian Rhythm of Water Intake in the Rat: Effects of Deuterium Oxide. <i>By</i> DIOMEDES E. LOGOTHETIS, ZIAD BOULOS, AND MICHAEL Terman.....	614
Self-Control and Responding during Reinforcement Delay. <i>By</i> TELMO E. PEÑA-CORREAL AND A. W. LOGUE	618
Time Discrimination versus Time Regulation: A Study on Cats. <i>By</i> F. MACAR, N. VITTON, AND J. REQUIN	622
The Effect of Variations in Reinforcement Probability on Preference in an Elapsed-Time Procedure. <i>By</i> CHARLOTTE MANDELL	625
Timing of Initiation and Termination of Dual Manual Movements. <i>By</i> MICHAEL PETERS	628
Explicit Counting and Time-Order Errors in Duration Discrimination. <i>By</i> WILLIAM M. PETRUSIC	630
Rats Remember the Circadian Phase of Feeding. <i>By</i> ALAN M. ROSENWASSER ...	634
Temporal Organization and Intermodality in Duration Discrimination of Short Empty Intervals. <i>By</i> ROBERT ROUSSEAU, JOCELYN POIRIER, AND GÉRARD TREMBLAY	636
An Adaptive Counter Model for Time Estimation. <i>By</i> ROBERT ROUSSEAU, DANIEL PICARD, AND EDGARD PITRE.....	639
The Decision Rule in Temporal Bisection. <i>By</i> STEPHEN F. SIEGEL AND RUSSELL M. CHURCH	643
Temporal Patterning and Selective Attention Effects on the Human Evoked Response. <i>By</i> JUNE J. SKELLY, ANTHONY RIZZUTO, AND GLENN WILSON ...	646
Chronological Knowledge: The Cognitive Integration of Temporal Information in Pigeons. <i>By</i> RON WEISMAN	649
A Computer-Based Portable Keyboard Monitor for Studying Timing Performance in Pianists. <i>By</i> CRAIG MINOR, MARK TODOROVICH, JAMES BOYK, AND GEORGE P. MOORE	651
Index of Contributors	653

Financial assistance was received from:

- INSTITUTE FOR COGNITIVE RESEARCH through COLUMBIA UNIVERSITY
- McMASTER UNIVERSITY
- NATIONAL INSTITUTE OF MENTAL HEALTH—NATIONAL INSTITUTES OF HEALTH
- NATIONAL SCIENCE FOUNDATION
- OFFICE OF NAVAL RESEARCH

The New York Academy of Sciences believes it has a responsibility to provide an open forum for discussion of scientific questions. The positions taken by the participants in the reported conferences are their own and not necessarily those of The Academy. The Academy has no intent to influence legislation by providing such forums.

PART I. TIME PERCEPTION

Introduction

JOHN GIBBON

*New York State Psychiatric Institute
New York, New York 10032; and
Department of Psychology
Columbia University
New York, New York 10027*

LORRAINE ALLAN

*Department of Psychology
McMaster University
Hamilton, Ontario, Canada L8S 4K1*

The study of timing and time perception has a venerable dual history in experimental psychology. Animal psychologists studying learning and conditioning have investigated extensively the timing and temporal patterning of behavior under instrumental control. In the psychophysical tradition, the psychophysics of time perception in humans also had early attention within the classical study of the limits of human sensory capacity. From these beginnings, the techniques and results in animal and human work have evolved and broadened so that currently there is a burgeoning literature in animal psychophysics dealing with time perception, and another in human timing dealing with the production and organization of temporal patterns.

The first part of this *Annal* exemplifies perhaps the greatest overlap of these two traditions. The papers in this section by Stubbs *et al.*, Eisler, Gibbon *et al.*, and Wasserman *et al.* present studies of time perception in rats and pigeons that bear quite directly on similar procedures or similar theories studied previously in the human. It is comforting to find that at least occasionally similar phenomena emerge. For example, Stubbs *et al.* report that pigeons, like humans, judge auditory durations as longer than visual durations. Further extensions of this trend, perhaps also including the study of human time perception with animal techniques, may be expected in the future.

Two papers are from the classical human psychophysical tradition and deal with perceived or subjective duration: Jamieson concentrates on providing a theoretical framework for the time-order error, and Allan describes two contingent duration aftereffects. She shows that these differ in important ways from the McCollough effect.

Another paper stems from the study of human information processing. This is Schweickert's analysis of critical path networks. His representations allow a determination of the temporal position of decision stages during processing.

Three of the papers develop models of time perception. They are good examples of the ways in which the two traditions are beginning to meld. Kristofferson's real-time criterion model for human duration discrimination postulates a minimal time "quantum" that determines all of the variance in the discriminative process. The model presented by Hopkins is similar in postulating central, deterministic (quantum) delays for response-stimulus synchronization performance in humans. The information-processing model developed by Gibbon *et al.* is applied to three different timing tasks performed by animal subjects. This account regards the procedures as differing in decision and memory processes, but utilizing the same central clock process. The pacemaker they propose for the internal clock bears structural similarities to the quantum process.

Quantal and Deterministic Timing in Human Duration Discrimination

ALFRED B. KRISTOFFERSON

*Department of Psychology
McMaster University
Hamilton, Ontario, Canada L8S 4K1*

When we began to study duration discrimination, we expected it to give us rather direct information about time perception. That expectation has not been fulfilled. Instead, our "thresholds" for duration seem to be determined by our ability to produce a time interval, to time it out internally. The size of a threshold is wholly determined by the extent to which repeated attempts to time out a fixed time interval are variable.¹

A set of duration stimuli is shown at the top of FIGURE 1. Each stimulus consists of two 10-msec auditory pulses separated by the stimulus duration, D . The stimuli differ from each other only in D , and the values of D are symmetrically arranged around a midpoint value. The midpoint of the set is the base duration. A single stimulus is presented on a trial and the subject is asked to categorize it as "long" (R_L) or "short" (R_S). Values of D greater than the midpoint are called long, and the decision on each trial is whether P_2 occurred before or after the midpoint value of D . In the experiments to be reported here, the subject is instructed to respond as quickly as possible, and the data consist of response probabilities and response latencies for each stimulus duration.

The general hypothesis is pictured in FIGURE 1B. On each trial, P_1 triggers an internally timed interval, I , which terminates as the criterion event, C . P_2 triggers a sensory event B_2 . If, as shown here, C occurs before B_2 , then a long response is determined. Short responses are triggered by B_2 whenever it occurs first. The discrimination mechanism is a race between the two response triggers, C and B_2 .

The two kinds of responses, therefore, have different causal histories, as shown in FIGURE 1C, R_L being linked to P_1 and R_S being linked to P_2 . Long responses should be time-locked to P_1 and should occur at the same time for all stimuli, that is, regardless of the time of occurrence of P_2 . R_S , on the other hand, should be time-locked to P_2 . These time-locking predictions are a major test of the hypothesis and they have been confirmed experimentally.¹

Since responding is speeded, short responses are direct reactions to P_2 and might resemble simple reaction times. Long responses are similar, except that the responses must be delayed, and R_L latencies might resemble time estimation latencies. These expectations are also confirmed, and speeded duration discrimination appears to be a combination of simple reaction time and time estimation, in which one kind of response occurs on a trial, the other being countermanded.

Practice with the stimulus set adjusts I so that C falls near the midpoint of the stimulus set. The lower panel in FIGURE 1 displays a specific model in which the times of occurrence of C are assumed to form an isosceles triangle. The variability in C is due solely to variance in I , the afferent latencies having zero variance. Therefore, for each D , B_2 is a fixed point which divides the triangle into two parts. The proportion of the area under the triangle to the left of B_2 represents the probability of R_L for that stimulus. Knowing the probability of R_L for two different stimuli, both of which have a B_2 within the triangle, enables one to calculate the quantum size, q , in msec, and also

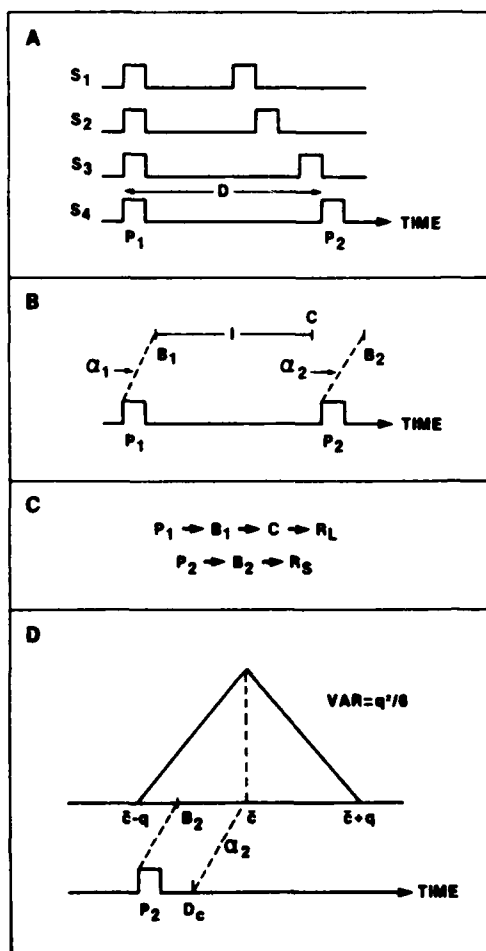


FIGURE 1. Schematic aids for the real-time criterion theory of duration discrimination. (A) a set of duration stimuli. Each pair of brief auditory pulses defines a stimulus duration, D . The midpoint of the set (MP) is halfway between D_2 and D_3 and is the base duration. (B) the general hypothesis. A race between C and B_2 determines which response will be made. (C) causal chains for the responses "long" and "short." (D) a specific quantal model. The times of occurrence of C are assumed to be distributed as an isosceles triangle having a base equal to twice the size of the time quantum.

D_C , the stimulus duration for which B_2 falls at the mean of the triangle. With a little practice, D_C is placed near the midpoint of the stimulus set.

We have shown in several experiments that the cumulative form of the triangle, which is a fully bounded sigmoid, is a satisfactory description of the psychometric function, at least as satisfactory as a normal ogive.²⁻⁴

The quantum size can also be calculated from the response latencies, giving a completely independent estimate of that parameter. This involves assuming that the two causal chains show in FIGURE 1 include the same variance sources except for the component C , an assumption already partially stated above. Consequently, the response latency variance for R_L is greater than that for R_S by an amount $q^2/6$.

The experimental tests that have been reported so far confirm all of the expectations described above.¹ Those experiments employed a single stimulus set, with a midpoint of 1150 msec, and one purpose of the experiments to be described now was to perform the same tests using other base durations. If the value of q is different at another base duration, then the magnitude of the change in q should be the same when it is calculated from response latencies as it is when calculated from response probabilities.

The value of q is a function of base duration,⁵ and the form of the function is embodied in a doubling rule: Doubling or halving base duration a given number of times, doubles or halves q the same number of times. This is illustrated in FIGURE 2.

In these experiments responding was not speeded, latencies were not measured, and the values of q were calculated from response probabilities only. Twenty consecutive sessions were conducted at one base duration before changing to another base duration. The dashed line shows performance during the first five sessions at each base duration, with the data points omitted here. Early in practice, q is directly proportional to base duration. Since the standard deviation of the discriminial process is $0.41 q$, the dashed line indicates that the ratio of S.D. to mean is constant at 0.053 during the first few sessions. This agrees closely with the ratio of 0.05–0.06 obtained by Getty.⁶

The data points and the line segments fitted to them in FIGURE 2 show performance during sessions 18–20. There is a practice effect which is specific to each base duration and which is large at some base durations and small at others. As a consequence of practice, steps unfold from the dashed line in such a way that the doubling rule is preserved. There appear to be steps at 200, 400, and 800 for this subject, with the quantum size doubling at each step.

However, if in the limit this is a step function, then that limit has not been reached in 20 sessions because the steps clearly slope upward slightly, and all by about the same amount. A second major purpose of the present experiments was to determine whether the steps become flat with even larger amounts of practice.

It is necessary to insert a methodologic note at this point. In measuring latencies, one finds two additional powerful sources of variance which complicate matters. One is competition between responses when two overt responses are used. We use a "go–no go" procedure in which the response may signal "long" and no response "short" in a particular experiment, or the reverse. This doubles the running time, but solves the problem. The second source of extraneous variance arises from properties of the response triggers C and B_2 . Each of them triggers its own response, and each also inhibits the opposite response. When C and B_2 occur in close succession, the first one determines the response outcome, but the second one perturbs the latency of that outcome. Hence, a minimum time separation must be assured between the two triggers in order to solve this problem, as will be shown.

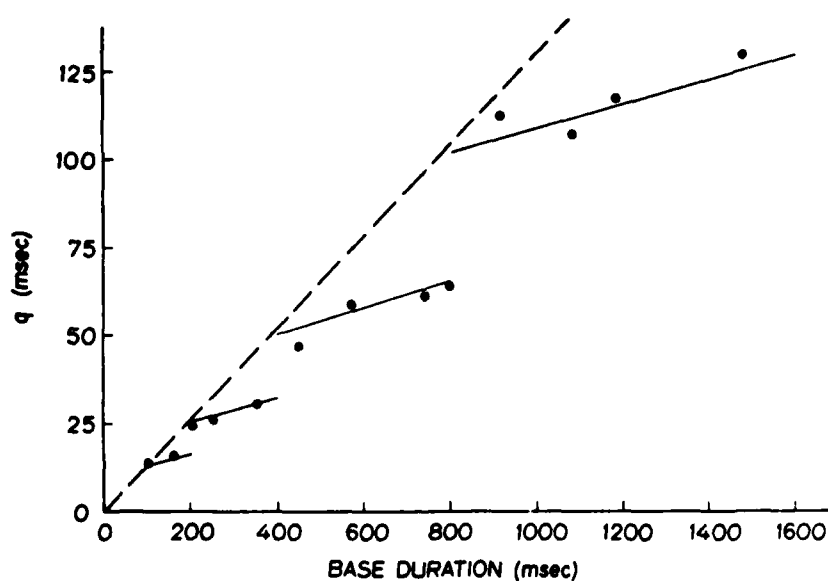


FIGURE 2. Quantum size as a function of base duration. The dashed line represents the results obtained with no prior practice. The points show the results that follow 17 sessions of practice at each base duration. (From Kristofferson.⁵ Reprinted by permission.)

Extensive data have been published that were obtained with me as the experimental subject.^{1,5} I am also the subject in the experiments to be described now, and it is possible to make quantitative predictions against which to compare these new data. In the 1977 experiment, the base duration was 1150 msec and speeded responding was used. The obtained value of q calculated from the response probabilities was 95.3. The measured latency variance for R_L was 1799 and that for R_S was 279 and the difference between these yields the second estimate of q , which is 95.5. Therefore, with an estimated $q = 95.4$, the two previously measured latency variances, the doubling rule, and the supposition that the steps are flat in the quantal step function, a basis is provided for calculating predicted values at other base durations.

When the base duration is 280 msec, as it is in the first three experiments, q should be on the 25-msec step and its predicted value is 23.9. FIGURE 3 shows the predicted distribution of C centered at 280 and with a base of 48 msec. The durations of S_2 and S_3 were fixed throughout, as indicated by the small arrows. In Experiment 1, the response was R_L and the duration of S_4 was varied over the indicated range, with S_1 fixed at 245. In Experiment 2, the response was R_S and S_1 was varied, with S_4 fixed at 305.

After sixty sessions of practice under the conditions of Experiment 1, 38 sessions were done with D_4 set at a different value from session to session so that each of the 19 values of D_4 were measured twice. For Experiment 2, the response was changed to R_S and 80 practice sessions were run with D_1 set at 255, five sessions at 250, and so on, down to 220. The parameters of the triangle calculated from the response probabilities actually obtained in Experiments 1 and 2 are also plotted in FIGURE 3. The span of the triangle is unaffected either by the variation in S_1 and S_4 or by the change in the meaning of the response. The centering of the triangle is slightly affected, but the total effect, amounting to only 1%, can be ignored for present purposes.

The predicted value of q is indicated by the dashed line in FIGURE 3. It agrees almost exactly with the mean of the obtained values, both being 23.9 msec.

The latency variances are shown in FIGURE 4. When S_1 or S_4 are close to the triangle, their latency variances are greatly inflated due to the temporal proximity of C

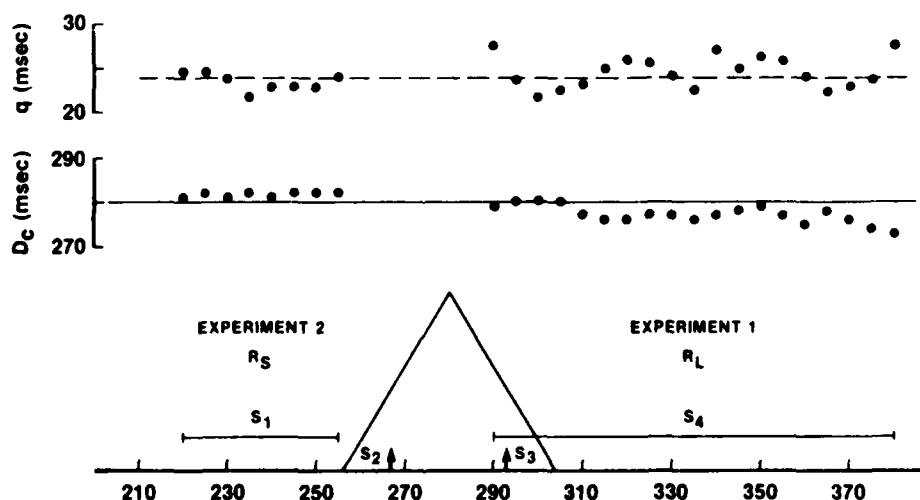


FIGURE 3. Experiments 1 and 2. The base duration was 280 msec and the triangle is the predicted distribution of C , centered at 280 and having a base of 48 msec. $D_2 = 267$, $D_3 = 293$. Experiment 1 varied only the duration of S_4 over the range shown. In Experiment 2, only D_1 was varied. The obtained values of D_C and q , calculated from the response proportions for S_2 and S_1 , are shown above. The dashed line is the predicted value of q .

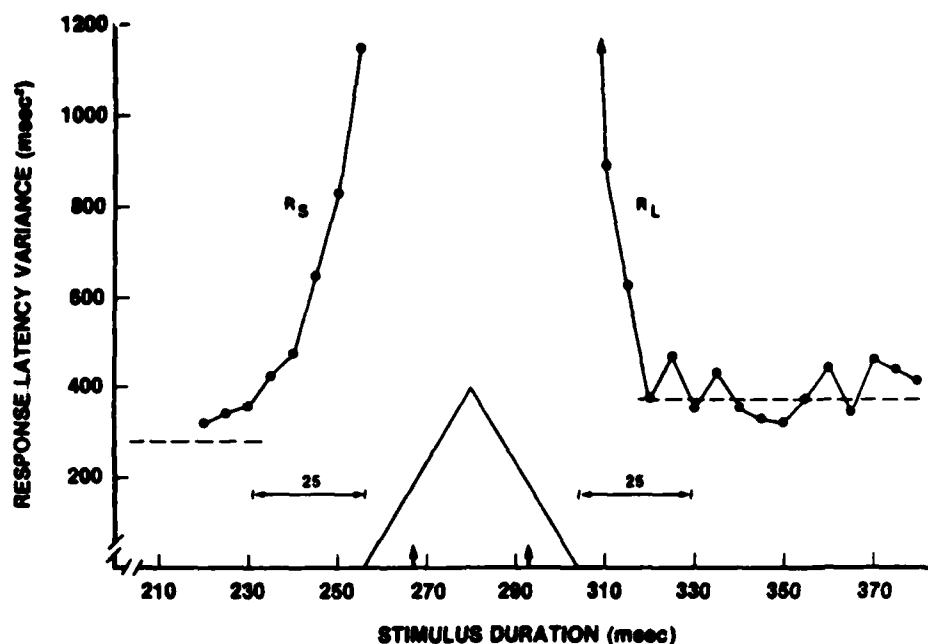


FIGURE 4. The same as FIGURE 3, except that the data plotted here are the response latency variances for R_S/S_1 on the left and R_L/S_4 on the right. The four values for $S_4 < 310$ are not shown; they are off the scale. As B_2 approaches the triangle from either direction, latency variance is greatly increased, but when B_2 is more than 25 msec from the nearest C , the latency variance is close to the predicted level, shown by the dashed lines.

and B_2 . The size of the effect is roughly the same above and below the triangle and it extends about 25 msec from the triangle in both directions. This suggests that the two response triggers, C and B_2 , must be separated by 25 msec or more to avoid latency interaction between them. It is important to add that within the two 25-msec windows, the latencies are affected, but response errors are not; the correct responses are given 100% of the time to stimuli within these windows.

The predicted latency variances are indicated by the dashed lines. The R_L variances were 1799 when the base duration was 1150; here, when the base is 280, they are much smaller, averaging 388 for the points more than 25 msec from the triangle. This agrees well with the predicted variance, which is 374.

The short response variance is affected very little by the change in base duration from 1150 to 280. It was very low at 1150,¹ and it appears to be slightly greater at 280. To determine its level more precisely, an additional 75 sessions were added to Experiment 2 with the duration of S_1 fixed throughout at a value that places it exactly 25 msec below the left corner of the triangle. The main results are presented in FIGURE 5. The value of q given by the response probabilities is stable and averages 23.8. There is a small, slow additional practice effect on the latency variance. After 35 sessions, the variance becomes extremely stable around a mean value of 280, the same as that reached when the base duration was 1150.

These experiments confirm the real-time criterion theory for short durations. Changing base duration from 1150 to 280 msec reduces the variance of the discriminial dispersion and the variance of the long-response latencies by the same large amount, an amount equal to that expected due to the change in quantum levels. The variance of the short-response latencies is unaffected. However, these results do not inform us about the flatness of the quantal steps. The next experiment was done for that purpose.

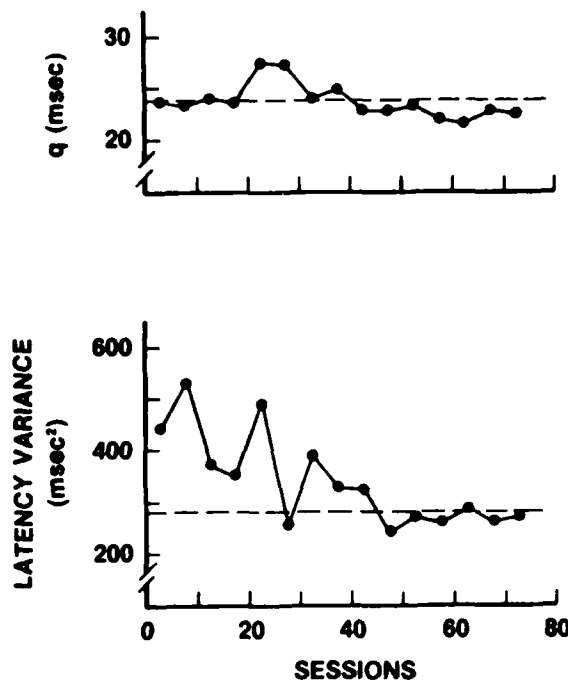


FIGURE 5. Experiment 3 (a continuation of Experiment 2 with all stimulus durations fixed for 75 sessions). D_4 was set at 305. D_1 was set at 235, which placed B_2 25 msec from the lower corner of the triangle when S_1 was presented. **Upper graph:** q calculated from response proportions for S_2 and S_1 . **Lower graph:** Latency variances for R_S/S_1 . The dashed lines are the predicted values.

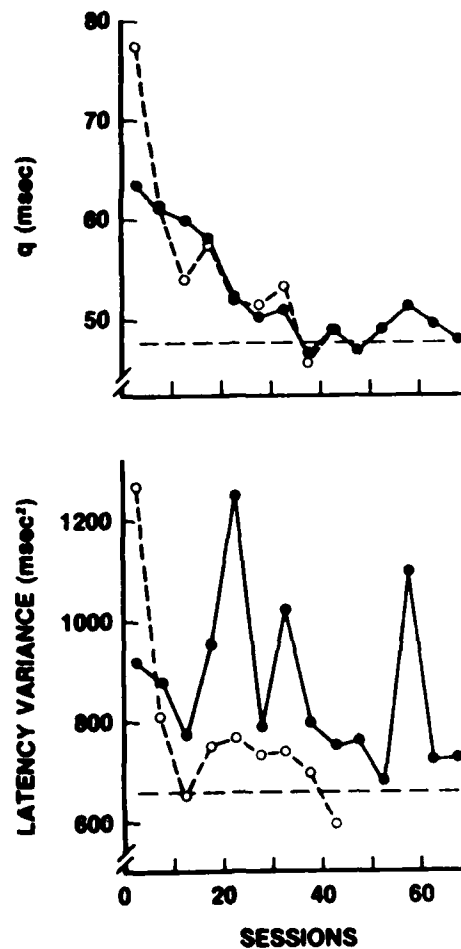


FIGURE 6. Experimental results at base durations of 560 (solid lines) and 750 msec (dashed lines) with $D_2 = 534$ and $D_1 = 586$ at 560 and $D_2 = 724$ and $D_3 = 776$ at 750. The horizontal dashed lines are the predicted levels, which are the same for the two base durations. **Upper graph:** Estimate of q based upon response probabilities; **lower graph:** latency variance of R_L on S_4 trials. As in all of the experiments reported here, errors are very rare when S_1 or S_4 are presented, $P(C)$ being 1.0 for S_4 and 0.999 for S_1 in these experiments.

Two new base durations, 560 and 750 msec, both selected from the q_{50} step, but widely separated on that step, were examined next. Seventy sessions were conducted at 560, followed by 45 sessions at 750 msec. The response was R_L throughout.

Once again, prolonged practice produced large changes in performance. The upper panel of FIGURE 6 shows that q , as obtained from the response probabilities, is the same for the two base durations after a few sessions of practice. It diminishes in value for about 35 sessions, thereafter averaging 48.6 at 560 and 47.3 at 750 msec, as compared to the predicted level of 47.7 indicated by the horizontal dashed line. The response latency variances (in the lower panel) are more erratic, but after 35 sessions they average 791 for the shorter base duration and 646 for the longer, compared to the predicted level of 658.

For the next experiment the base duration was 1500 msec, a duration that is both

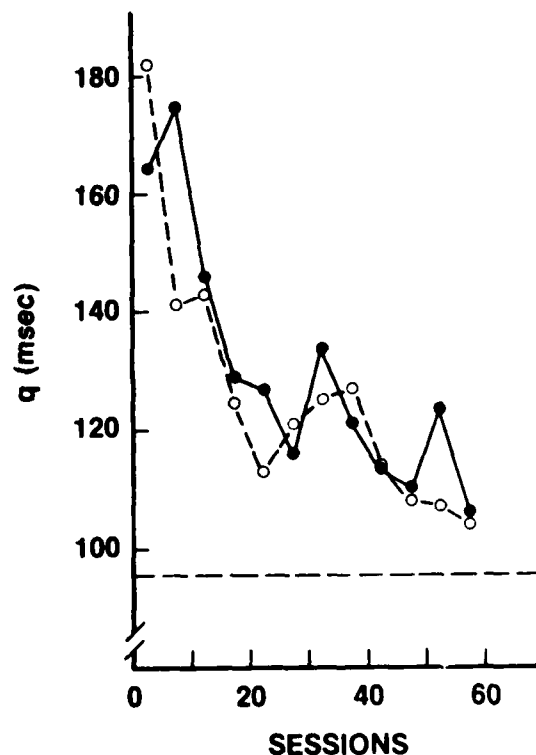


FIGURE 7. Results for 60 consecutive sessions at a base duration of 1500 msec. *Solid line:* Estimate of q calculated from response probabilities for S_2 and S_3 ; *dashed line:* estimates calculated from latency variance of R_L on S_4 trials, using 279 msec^2 as the latency variance of R_5 (see text). The *horizontal dashed line* is the predicted level. $D_2 = 1448$; $D_3 = 1552$.

long and is also situated near the high end of its quantal step, in this case the q_{100} step. A large practice gain is needed to unfold this location on the step from the Weber's law line.

The results of a 60-session series at 1500 are presented in FIGURE 7. The solid line shows the response probability, q . The latency variance for S_4 has been transformed into its associated value of q and that is indicated by the dashed curve. The practice effect is large and the two independent measurements of q are very similar throughout practice. This suggests that the practice effect is solely a reduction in the variability of I .

While the practice gain is large and relatively rapid during the first 25 sessions, it is obviously not complete after the usual 35 sessions, or for that matter even by session 60. The predicted level is approached, but not attained. During the final five sessions, the error of prediction is 11%.

The final experiment in this series had two purposes. There is the obvious question about the existence of a q_{200} level and a step up to that level near a base duration of 1600. To find out, a base duration of 1800 msec was studied next. This base duration would be close to the lower boundary of the step, if one exists, and the practice effect should be minimal for that reason, even though the base duration is even longer than the 1500 of FIGURE 7.

Twenty sessions were sufficient to answer both questions in the affirmative, as FIGURE 8 demonstrates. Unlike previous figures, FIGURE 8 plots single-session values. The solid circles show q as calculated from response probabilities, and the open dots indicate q as calculated from the latency variance.

A stable limit is achieved after only two sessions of practice. The line in the figure is fitted to all of the data points after session 2, and its slope is 0.077. The two measurements of q agree closely, both averaging 204.

This result suggests that the doubles set of q should be expanded to include q_{200} and that the quantal principle need not be confined only to very brief time periods.

The doubling hypothesis predictions discussed above have fared quite well for the various base durations that have been investigated. Taken together, the results show that in the limit, the quantal steps of the base duration function do become flat. This is summarized in FIGURE 9, where the results of the various base durations, including 1150,¹ are brought together on a single step, the q_{50} step, by applying the doubling rule. For example, the points plotted here at the 450-base duration are from the base duration of 1800, with both 1800 and the obtained values of q halved twice.

Only those experiments in which the response was R_L , and in which D_4 placed B_2 more than 25 msec above the triangle, are included. In all cases, the amount of prior practice at the base duration is large, at least 35 sessions, except for the 1800-base duration, as pointed out above.

The solid circles show the value of q_{50} calculated from the response probabilities for S_2 and S_3 . The open circles show the values calculated from the R_L latency variance for

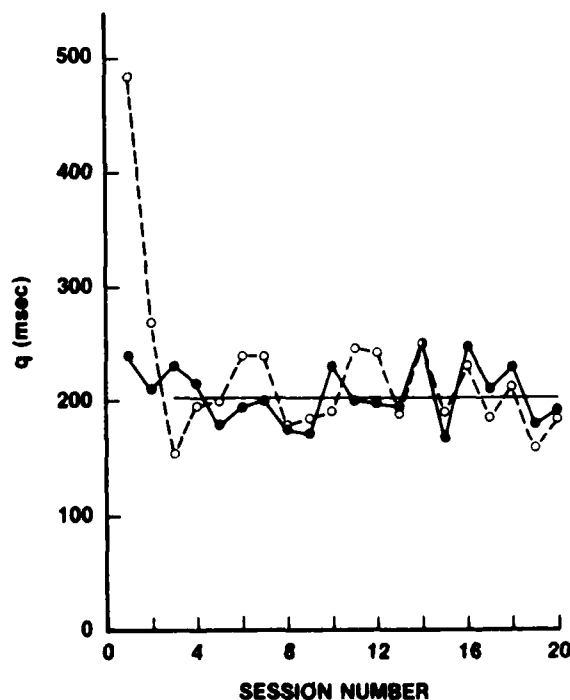


FIGURE 8. Twenty sessions at a base duration of 1800 msec. *Solid circles:* q calculated from response probabilities for S_2 and S_3 ; *open circles:* q calculated from response latency variance on S_4 trials. The *solid line* is the least-squares line fitted to all of the data after session 2; it has a slope of +0.077. $D_2 = 1700$; $D_3 = 1900$.

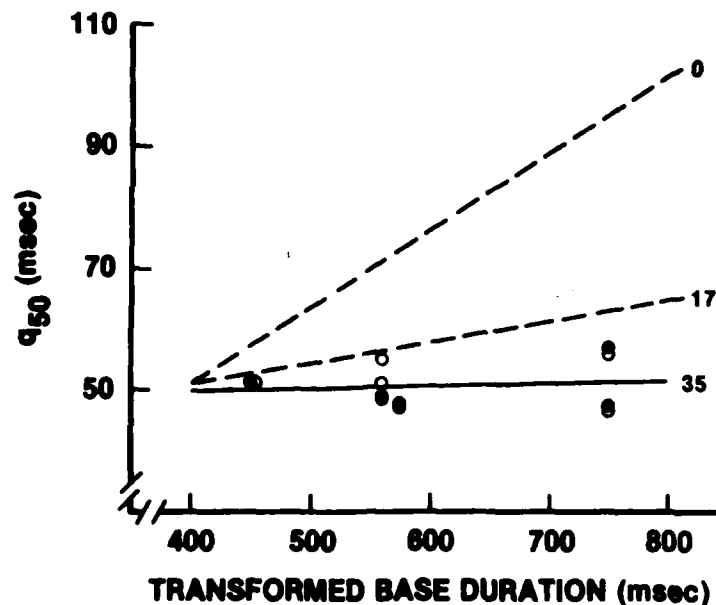


FIGURE 9. Summary of all experiments. The data points show the best estimate of q after at least 35 sessions of practice. *Solid circles:* q from response probabilities; *open circles:* q from latency variances. All base durations are included by doubling or halving both q and base duration so as to bring all within the Q_{50} range. The *solid line* is fitted to all of the data and has a slope of almost zero. The *upper dashed line* shows performance with no practice and the *lower dashed line* shows performance with 17 practice sessions (both from FIGURE 2).

S_4 , using 279 as the estimate of R_5 latency variance. These two measurements of q agree closely, confirming the real-time criterion theory over the full range of base durations.

The solid line is fitted to all of the data points and its slope is close to zero, indicating that the steps do become flat with sufficient practice. The upper dashed line shows the result obtained with little or no practice and the middle line the result obtained after 17 sessions of practice.⁵ With increasing practice, the steps unfold progressively and finally become flat. The hinge is very close to 400 msec. These experiments were done at various times during a 10-year period and the mechanisms that are involved appear to be quite stable.

DISCUSSION

When the time quantum concept was proposed in 1967, the supporting experiments included several different ways to measure the quantum size, each of which gave a value near 50 msec.⁷ Later experimental work gradually revealed that if the concept were to be retained with some degree of generality, values other than 50 would have to be accepted. For example, under certain conditions, successiveness discrimination functions appeared to require a mixture of quantum sizes of 50 and 100.^{8,12} Also, our initial work on duration discrimination gave us values of 25 as well as 50 (Ref. 2). There were other such instances.

The step function in duration discrimination described above is the first single function to reveal a full range of quantum sizes. The permissible values form a doubles

set extending from 12 to 200 msec. The two end values of this set are not yet firmly established by multiple operations and their status is tentative.

The step function has a second important implication. Since the steps have flat trends, deterministic timing appears to be involved. For example, for all base durations between 200 and 400 msec, $q = 25$. Changing the base duration changes the mean value of the internally timed interval, I (see FIGURE 1). Since q expresses the variability of I , the conclusion is that the mean of I can be changed without affecting its variance. Within limits, the timing of I is deterministic. When $q = 25$, deterministic timing up to at least 400 msec is possible.

This is our second report of deterministic timing. The first appeared several years

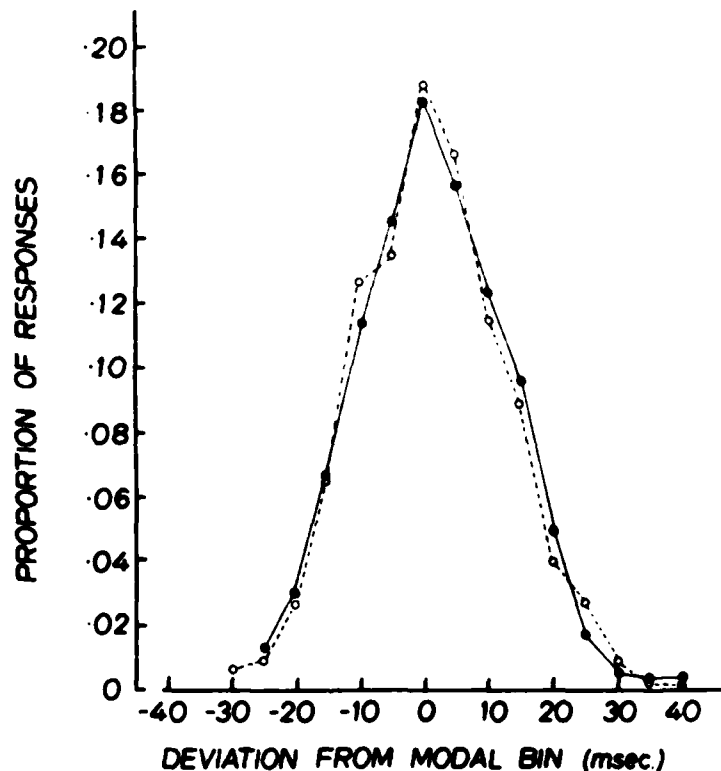


FIGURE 10. Two distributions of response latencies in response-stimulus synchronization. *Solid line:* mean = 547, *S.D.* = 12, *N* = 600; *dashed line:* mean = 307, *S.D.* = 12, *N* = 900. (From Kristofferson.⁹ Reprinted by permission.)

ago in experiments on response-stimulus synchronization.⁹ There too, the upper limit was found to be 400 when q is 25. Sample data are shown in FIGURE 10.

Two distributions of latencies for a response initiated by a brief auditory pulse are plotted in FIGURE 10. The two distributions appear to be identical, but actually they differ in one respect: one has a mean of 300 msec, while the mean of the other is 550 msec. The mean latency can be as short as the reaction time limit (here, about 160 msec), or the response can be delayed some additional time, up to 400 msec, without changing the variance. This internally timed interval can be set at any value between zero and 400 and give this result.

The quantum size here is 25, and the distributions are the convolution of an

isosceles triangle having a base of 50, with a low variance, efferent delay distribution having a standard deviation of about 6 msec. A separate means of assessing the efferent delay variance, which agrees with this estimate, is provided by the theory and data developed by Alan Wing concerning interresponse timing.¹⁰

These distributions describe the latencies of an element S-R chain. They have variance less than that of simple reaction times by a factor of 4 or more.¹¹ They have been obtained for quantum sizes of 50 (Ref. 9), 25 (shown above), and 12 (Ref. 11). The last value provides some additional evidence for a 12-msec quantum size. Gordon Hopkins will pursue this in greater depth, including the convolution interpretation mentioned above, in this volume.

The duration discrimination step function extends our description of deterministic timing by showing that the upper limit of deterministic timing is not merely the single value of 400 msec. Instead, the upper limit changes, depending upon the value of q that is in effect. For example, when q is 12, the upper limit is 100, and when q is 100, deterministic timing occurs out to 1600 msec. On the basis of these data, the rule seems to be that the upper limit of deterministic timing is 16 times the current quantum size. This encourages the view that quantal timing and deterministic timing are not the result of completely separate mechanisms. However, I resist the temptation to propose a mechanism for two reasons. The first reason is obvious: extensive data for the step function are available so far only for the single subject discussed above. The second reason is that, while the 16- q rule did also hold in the initial synchronization experiments, it has also failed. Hopkins and Kristofferson¹¹ found deterministic timing to an upper limit of at least 400 with q in the 12-msec range.

There is another reason to be cautious when generalizing from the data presented above. The difficulty level of the discrimination is determined by the size of ΔD , the difference in duration between S_2 and S_3 . While I am confident that the triangle is a good description of the discriminial dispersion, and therefore that q is independent of ΔD for a given distribution of C , I cannot say that the value of q attained asymptotically with practice is independent of ΔD . When ΔD is fixed throughout practice, as it was here, the asymptotic q might possibly be affected by the value of ΔD that is used. This possibility must be investigated.

The quantum size, therefore, can no longer be considered a single value. Instead, it may take on any value from the doubles set. There is no compelling evidence that it is variable beyond that, however. I point this out because some of the graphs presented herein, such as those that show the apparent gradual decrease in q with practice, may be taken to imply that q varies continuously. The data were presented this way for ease of exposition, with q as calculated for preasymptotic sessions meant to be only a rough estimate of variability. Actually, we do not yet have anything more than guesses concerning the reasons for the variance reductions observed to occur as a result of practice.

A few empirical principles of timing are beginning to take shape. My own list has only three entries, followed by many blank spaces. Two are the quantal and the deterministic principles that I have been discussing, and there is also a statistical principle, embodied in Weber's law, which I have only mentioned, but for which one can find a substantial amount of support.

On the level of theory we are still quite unconstrained so far as general directions are concerned. My own preference is to postulate a periodic process as a generator of quantal variability and to use the deterministic nature of the periodic process to explain deterministic timing, thereby linking the first two principles. Since the standard deviation of some quantal distributions is directly proportional to quantum size, the third principle, Weber's law, could also be brought within such a theory by postulating continuously variable quanta in some parts of the system.

Other theoretical directions are possible, of course. To take an extreme example, an essentially probabilistic model could be constructed to explain deterministic timing. One could postulate that the variability of the added delays does increase as the mean increases but that there is also a gradient of increasingly negative correlation which just balances the variance increase. That does not seem plausible to me, but it is possible. It would be difficult to extend such a model to explain quantal doubling and the doubling of the upper bound on deterministic timing.

It is obvious that the mere fact that a system displays temporal regularities does not mean that the system contains dedicated timing mechanisms. As far as we know, the system with which we deal may contain any number of clocks, including none at all. The empirical principles that we adduce might all describe nothing more than temporal properties of the system. Whether some subset of the principles reflects the operation of a clock or clocks, we do not know.

The proposition that time quanta are generated by a periodic process does not imply the existence of a clock. The hypothesis that there is exactly one clock that controls timing at all loci in the system seems to me to be the most likely candidate for rejection. One reason is that we repeatedly find that quantal delays at different loci fluctuate independently of each other. (For example, the triangular distribution is the distribution of the sums of two identically distributed quantal delays, each uniformly distributed, only if the two quantal delays are statistically independent). How could that be so if a single clock were controlling both loci?

Another reason is to be found in the analysis of the response "long" in the real-time criterion theory. There appear to be two pairs of quantal loci in the chain between the stimulus (P_1) and R_L (see FIGURE 1). One pair precedes C and the other follows C . Changing base duration can change the quantum level of the first pair, but does not change that of the second pair. A single clock would have to be able to generate two frequencies "almost" simultaneously. To strengthen this argument, we need to find out whether two real-time criteria can be timed during overlapping periods of time and at different quantum levels.

I believe we are being pushed in the direction of having to postulate multiple clocks. The greater the number, the less point there seems to be to talk of dedicated mechanisms, and we will be left where perhaps we should agree we now are: talking about temporal properties of the system.

REFERENCES

1. KRISTOFFERSON, A. B. 1977. A real-time criterion theory of duration discrimination. *Percept. Psychophys.* 21(2): 105-117.
2. ALLAN, L. G., A. B. KRISTOFFERSON & E. W. WIENS. 1971. Duration discrimination of brief light flashes. *Percept. Psychophys.* 9: 327-334.
3. ALLAN, L. G. & A. B. KRISTOFFERSON. 1974. Judgments about the duration of brief stimuli. *Percept. Psychophys.* 15: 434-440.
4. KRISTOFFERSON, A. B. & L. G. ALLAN. 1973. Successiveness and duration discrimination. *In* Attention and Performance: IV. S. Kornblum, Ed. Academic Press, New York, NY.
5. KRISTOFFERSON, A. B. 1980. A quantal step function in duration discrimination. *Percept. Psychophys.* 27(4): 300-306.
6. GETTY, D. J. 1975. Discrimination of short temporal intervals: A comparison of two models. *Percept. Psychophys.* 18: 1-8.
7. KRISTOFFERSON, A. B. 1967. Attention and psychophysical time. *Acta Psychol.* 27: 93-100.
8. KRISTOFFERSON, A. B. 1967. Successiveness discrimination as a two-state quantal process. *Science* 158: 1337-1339.

9. KRISTOFFERSON, A. B. 1976. Low-variance stimulus-response latencies: Deterministic internal delays? *Percept. Psychophys.* 20(2): 89-100.
10. WING, A. M. & A. B. KRISTOFFERSON. 1973. Response delays and the timing of discrete motor responses. *Percept. Psychophys.* 14(1): 5-12.
11. HOPKINS, G. W. & A. B. KRISTOFFERSON. 1980. Ultrastable stimulus-response latencies: Acquisition and stimulus control. *Percept. Psychophys.* 27(3): 241-250.
12. ALLAN, L. G. & A. B. KRISTOFFERSON. 1974. Successiveness discrimination: Two models. *Percept. Psychophys.* 15: 37-46.

Ultrastable Stimulus-Response Latencies: Towards a Model of Response-Stimulus Synchronization

GORDON W. HOPKINS^a

*Department of Psychology
University of Alberta
Edmonton, Alberta, Canada T6G 2E9*

Over the past few years I have been investigating the remarkable ability of human beings to accurately anticipate the time of occurrence of a predictable sensory event and to synchronize an overt response to that event. Presumably, this type of response-stimulus synchronization behavior is mediated by central temporal mechanisms which time the delay required to trigger the response such that it occurs temporally coincident with stimulus onset. In our examination of the nature and functioning of these human temporal mechanisms, a major aim was to develop special procedures for minimizing response latency variances to facilitate mathematical simulation and modeling of the information-processing stages involved in this type of stimulus-response chain.

The basic task, modeled after Kristofferson,¹ is diagrammed in FIGURE 1. On each trial, contact of the subject's index finger with a small touch-sensitive switch (R_1) generated a short, variable foreperiod followed by presentation of two brief auditory stimuli, P_1 and P_2 , separated by a short time interval, referred to as the P_1P_2 interval. This interstimulus interval was fixed for each subject, but varied across subjects from 310 to 550 msec. The subjects were instructed to anticipate the second stimulus, timing from the first (P_1), in order to trigger a finger withdrawal or synchronization response (R_2) which would be manifested in synchrony with onset of the second stimulus (P_2). Thus, response latency was measured from P_1 onset to the moment when the subject's finger broke electrical contact with the response button. Immediate perceptual feedback regarding the accuracy of response was available to the subject by attending to the temporal order relationship between R_2 and the onset of P_2 . In addition, a delayed feedback signal was provided, in the form of a third auditory pulse, which had a duration equal to the error of synchrony, the time difference between R_2 and P_2 onset. The direction of error was signaled by a light accompanying the delayed feedback pulse when the error was positive, that is, when R_2 occurred after P_2 onset.

This procedure that I have outlined institutes several modifications of the basic response-stimulus synchronization procedure used by Kristofferson¹; these include the use of subject-paced trials and provision of highly salient feedback. Also, a special technique was developed that involved manipulation of the foreperiod between R_1 and P_1 onset in conjunction with independent deletion of each of the stimulus components in a trial sequence. This kind of control allowed unambiguous determination of the role played by each of the stimuli in maintaining synchronization performance.

These modifications resulted in a significant reduction in the lowest previous estimate of response latency variance. Minimum variances under 35 msec² were

^aPresent address: Communications Research Centre, P.O. Box 11490, Station "H," Ottawa, Ontario, Canada K2H 852.

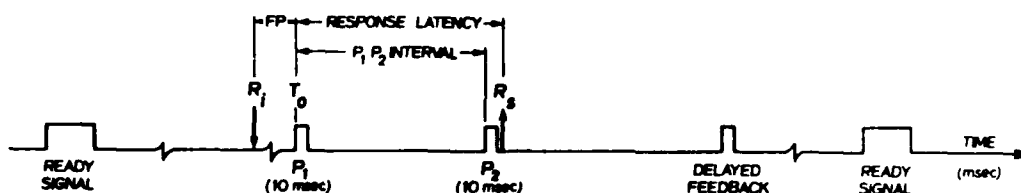


FIGURE 1. Diagram of a typical synchronization trial. Spacing of trials is paced by the subject's initiation response, R_i , which can be made at any time following ready signal. FP refers to the foreperiod duration.

obtained and the data indicated that response latency variance was independent of mean latency over the range of synchronization intervals from 310 to 550 msec. Within this range, latency distributions were the same, symmetrical, and sharp-peaked, with all responses contained within a 50-msec time window. A typical response latency distribution exhibiting these characteristics is shown in FIGURE 2. This relative frequency distribution combines 1500 response latencies obtained at a P_1P_2 interval of 460 msec and is plotted using a bin size of 3 msec. The mean is 461 msec with an overall variance of 35 msec.² The role of feedback in accurate synchronization performance was also examined, using the manipulations outlined previously. This provided data that indicated feedback to be one of the most important factors responsible for producing and maintaining the ultrastable, low-variance stimulus-response latencies observed. The other important factor appears to be prolonged practice at a particular synchronization interval.

Results of these studies all provide support for Kristofferson's^{1,3} notion of nonvariable, centrally timed delays which can be inserted into a stimulus-response

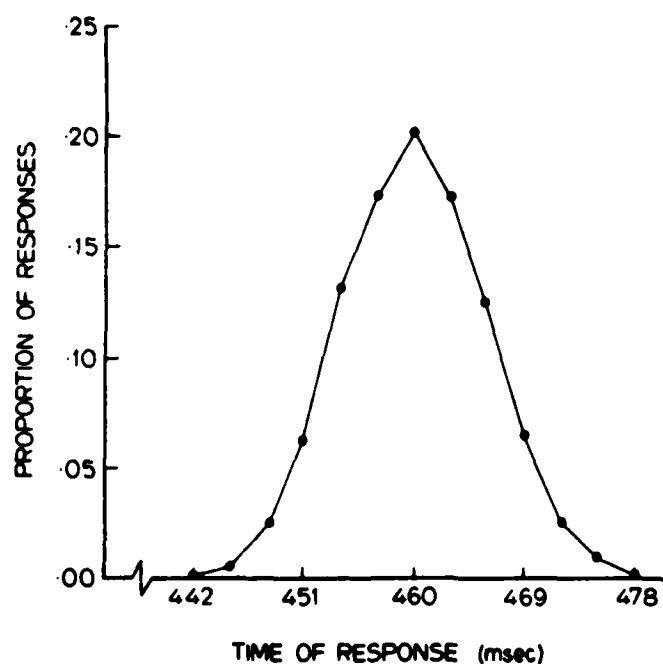


FIGURE 2. Low-variance, relative frequency distribution containing 1500 response latencies for subject G.H. on sessions 266-270 with a P_1P_2 interval of 460 msec (mean = 461, total variance 35, mean within-block variance = 34, variable foreperiod). Bin size is 3 msec.

chain. These delays are easily adjustable, but once set, are deterministic. The aim of this presentation is to examine this notion in the response-stimulus synchronization context, to formulate a mathematical model on the basis of the data obtained from carefully controlled experiments, and to try to develop a theoretical framework for conceptualizing this type of behavior.

Several models have been proposed to account for the various timing capacities of the human central nervous system. Many of the mechanisms incorporated into these models hypothesize a "time base" of some sort which generates a succession of temporal cues that can be used by the central nervous system for response timing⁴ and for controlling the gating of information flow from one central stage to another.⁵

In some models the internal clock varies somewhat in rate, causing successive temporal judgments to be variable. The mechanism for timing in these models involves accumulating clock pulses, during the duration to be judged, from a source with identically distributed interpulse delays. In Creelman's⁶ model this source is assumed to be Poisson-distributed, whereas for Triesman,⁷ the nature of the distribution is not specified. In both cases, however, the models predict increasing variance in temporal judgments as a function of the mean interval to be represented. In some experiments the results suggest an increasing linear function between variance and mean,⁶ whereas other results suggest a similar relationship, but between standard deviation and mean.^{8,9} In any event, neither model is appropriate for describing the current response-stimulus synchronization data in which variance and mean response latency are independent over a substantial range of temporal intervals, suggesting that some kind of deterministic timing mechanism might be involved.

Kristofferson¹ proposed a deterministic type of model to account for the response-stimulus synchronization performance he observed in his experiments. Latency distributions revealed a simple, homogeneous stimulus-response unit that was the same whether the mean was 160 msec or 550 msec. Kristofferson proposed that the elementary response latency distributions observed resulted from the convolution of three independent sources of variance inherent in the stimulus-response chain.

One of these component distributions is normally distributed and represents variability in the efferent delay between the time when the response is triggered internally and when the overt response is produced. Afferent latencies, on the other hand, are assumed deterministic or nonvariable. Support for this assumption comes from temporal order discrimination¹⁰ and duration discrimination data.²

The other two sources of variance are assumed to be identical and independent uniform distributions spanning a range of one time quantum. When convoluted, these produce a triangular distribution spanning two time quantum units. These delays, according to the model, represent variable delays in the processing of the stimulus-response chain, but the exact nature or locus of these delays is left unspecified. Also, there is no mechanism proposed to account for the assumption of independence between the quantal units. If these delays result from the operation of a single central mechanism, which gates information through the central information processor, then it is difficult to explain how the two delays, assumed quantal in nature, and their associated variances can be considered independent because they presumably are both dependent upon the same nonrandom underlying process.

Despite these criticisms, Kristofferson's model does provide quite accurate predictions of the asymptotic response latency variance obtained using the modified response-stimulus synchronization procedure described earlier. These predictions were based on an estimate of a minimum time quantum of 12 msec, suggested by some^{2,3} duration discrimination work, and a minimum efferent delay variance of 10 msec,² based on interresponse interval timing experiments.¹¹ The model, however, was never tested for its "goodness-of-fit" to the data. Thus, the aim of the rest of this

discussion is to further examine new empirical data relevant to specifying characteristics of a revised model of anticipatory timing, to outline the resulting model in detail, and to provide mathematical support for the model's ability to accurately represent the data.

In response-stimulus synchronization, variability on the temporal axis can arise from several possible sources. These include inconsistency in the afferent delay between P_1 onset and its central registration, variability associated with the timekeeping process itself, and variance in the efferent delay between response trigger and overt response. However, as mentioned earlier, data from several experiments suggest that afferent latencies in this task can be considered deterministic, which means that although the afferent delay has some non-zero value associated with it, the variance is negligible. Thus, only variability in central timing and output of the response need be considered.

In the model being developed, the central and motor components are assumed, on the basis of several pieces of evidence, to be independent. For example, an analysis of carefully collected interresponse timing data conducted by Wing¹¹ indicated that response latency variance was basically a constant and was independent of mean interresponse interval. Total variance increased with interresponse interval, but this was attributed to increases in variability of the central processing component responsible for triggering the overt responses.

Further support comes from several simple and delayed reaction time (RT) experiments, in which electromyographs (EMGs) were taken while response times were measured, after which calculations were made of the correlation between pre-motor time (time from onset of the action stimulus to EMG onset) and response time versus the correlation between motor time (time from EMG onset to the overt response) and response time. The results revealed correlation coefficients close to zero between motor time and response time, indicating independence between central and motor components.¹² Also, in the delayed RT situation, the EMG activation preceded the overt response by a relatively constant interval regardless of the actual response latency produced.¹³

A similar result was obtained by Michaels¹⁴ using a response-stimulus synchronization paradigm and a countermanding procedure. The subject's task was to withhold the synchronization response if a third signal occurred during the P_1P_2 interval. The data were proportions of correctly countermanded responses as a function of the time between the countermand signal and P_2 onset. From this data an estimate of response trigger timing was derived and the results indicate that the trigger always precedes P_2 onset by a fixed time period, independent of the P_1P_2 interval.

These findings lead to the conclusion that manipulations involving anticipatory response affect only the central, pre-motor component of a stimulus-response chain and argue against efferent stages' having any major participation in timekeeping, at least in this rather simplified paradigm. Obviously, for extended chains of motor behavior, timing may become more complex. Tyldesley and Whiting¹⁵ among others¹⁶⁻¹⁹ have suggested that in these situations some "timing" is simply a byproduct of efferent delays generated by each motor component. However, this discussion will not address that type of timing.

The last piece of evidence in support of the notion of independence between central and efferent stages comes directly from our synchronization experiments. Findings of independence between response latency variance and mean over a wide range of P_1P_2 intervals strongly suggest the existence of an adjustable, nonvariable, central delay mechanism combined with a constant mean efferent latency. This is so because if response latency was lengthened by increasing the number of motor components in the S-R chain, each with its own inherent variability, then there should be a commensu-

rate increase in variance. But this is not the case. Consequently, we assume that efferent delays are distributed with constant mean and standard deviation despite relatively large changes in the overall mean response latencies produced during synchronization at different P_1P_2 intervals.

The last point to be discussed, before formulation of the model, concerns the accuracy and variance associated with the motor component of response latencies. The overt response is simply a finger withdrawal, but actually several motor elements are involved in producing this movement. The interesting point is that although each of the underlying elements, when measured separately, exhibits a rather large temporal variation, the outcome of their joint action produces a response well defined in time. Meijers and Eijkman²⁰ have examined this apparent paradox and offer an explanation based on the macro-activity required for the motor system to initiate elementary movements. They show that execution of an overt response requires the joint effort of many elements and it is this requirement of joint activity that allows the remarkably small stochastic variation observed to be obtained. This is accomplished by summation of element activities, thereby providing a better time definition than that produced by any of the individual activities. In other words, averaging the behavior of several elements can cancel the effect of individual temporal inaccuracies.²⁰

From the foregoing discussion, it is obvious that several factors must be considered in formulating a mathematical model that will not only provide a good representation of the data, but will also have parameters that are psychologically relevant. The latter stipulation is important because without it the model will have little utility in generating the testable predictions needed for furthering our understanding of the internal mechanisms that underlie response-stimulus synchronization behavior, thereby extending existing theory.

THE MODEL

Consequently, a rather traditional approach was taken in formulating the model. The approach is based on the premise that if the data are stable, then any stochastic processes associated with delays accrued in each of a series of independent processing stages would be reflected in an overall distribution of response latencies given by the convolution of all of the component distributions.

The response-stimulus synchronization model that is being developed is similar to Kristofferson's,¹ but includes some modifications and extensions in an attempt to provide a locus for the central delays incurred in processing the stimulus-response chain as well as to account for the assumption of independence of the central, stochastic components. In describing the model, I will employ some computer metaphors to facilitate an understanding of the mechanisms involved. However, the use of these metaphors is purely for eclectic reasons and should not be construed as implying any direct analogy. I would also like to note that the first part of this discussion assumes that a steady-state condition exists in the central information processor. Violation of this assumption will be dealt with subsequently.

FIGURE 3 diagrams the modified response-stimulus synchronization model. The onset of P_1 is a sharply defined external event, although its sensory effect is extended over time, as shown by the interval labeled *afferent latency*. This latency refers to the time from onset of the peripheral stimulus until an internal state has developed, as a result of P_1 stimulation, that is sufficient to exceed some criterion and trigger the next stage in the information-processing chain. This process is similar to filling an input buffer and setting a flag that indicates that information is available for further

processing. In this context, the afferent latency is a combination of the transduction time at the peripheral receptor, the conduction time from periphery to the central system, plus the time required to represent this information in one of the registers of the central processing unit.

Although the afferent latency certainly has some non-zero value, its variability is assumed to be negligible on the basis of several pieces of evidence presented earlier. Therefore, afferent latency can be considered a constant, contributing nothing to the variance and shape of the observed distribution of response latencies.

Once the stimulation produced by P_1 onset is registered centrally, in a buffer, the information must wait for a period represented by W_1 before gaining access to subsequent processing stages. This waiting time for information transfer from input buffer to the deterministic timekeeping mechanism results because the contents of the

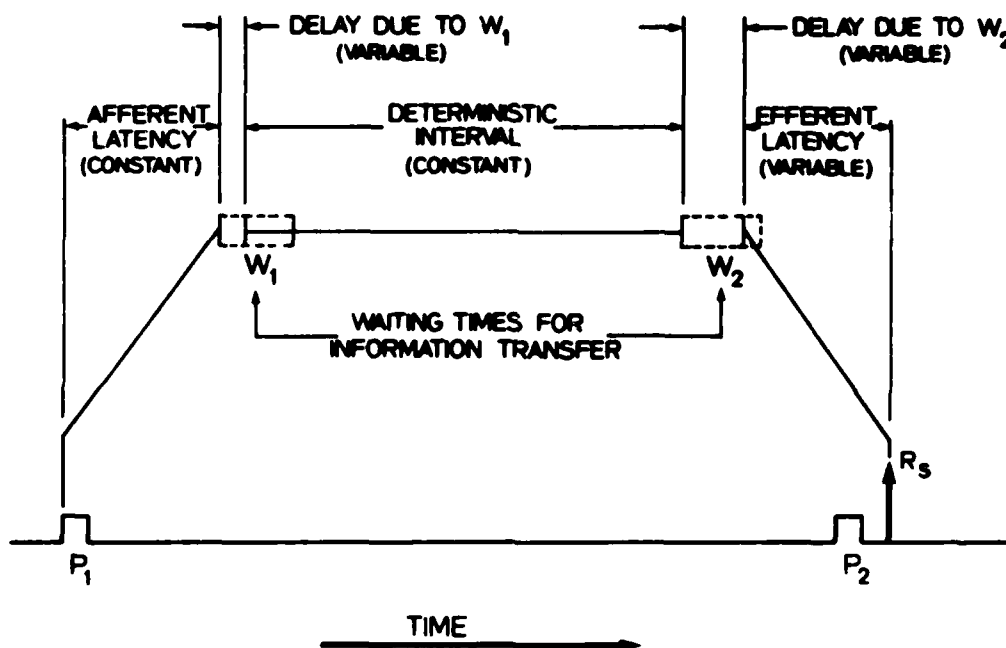


FIGURE 3. Diagram of response-stimulus synchronization model. P_1 , P_2 , and R_s are external, observable events, while W_1 and W_2 are hypothesized internal events. The diagram represents the time course of various components of a stimulus-response chain on a typical trial.

input buffer are only accessed periodically. Every n units of time the central processor reads the contents of the input buffer and performs operations based on this information. Because this cycle time, or scheduling of access time-points, is independent of peripheral stimulation, the delay due to W_1 is variable and uniformly distributed over a range from zero to W_{1max} milliseconds. For example, sometimes stimulus information will reach the input buffer just prior to the start of a new cycle and thus will gain access to the central processor with very little delay. On the other hand, stimulus information loaded into the input buffer just after the start of a cycle will have to wait almost an entire cycle, W_{1max} , before gaining access to the central processor.

After access of stimulus information to the central processor, or in this case, the deterministic timekeeper mechanism, a delay appropriate to the synchronization

interval and state of the organism is assumed to be generated. Since this delay has negligible variance associated with it, the term deterministic interval is used. Physiological mechanisms capable of producing such delays are not forthcoming from the literature, but the rationale for assuming the existence of such mechanisms is clear from the preceding discussion and experiments.

The termination of this deterministic delay produces information that is then loaded into an output buffer and generates a flag in a fashion similar to that described for the input buffer. This time, however, the information is waiting to gain access to the response processor. This processor also has a fixed cycle time, W_{2max} , which is similar to that of the central processor, W_{1max} , thus generating uniformly distributed delays over a range from zero to W_{2max} . Although W_{1max} and W_{2max} are similar, they are not exactly the same. As a result the two processors cycle in and out of phase relatively frequently during the synchronization interval. Consequently, the two waiting times can be considered independent since the initiation of a synchronization interval (P_1 onset) is totally independent of any phase relationship existing between central and response processor cycle times.

An example may help to clarify this mechanism that I am proposing to allow the assumption of independence between W_1 and W_2 . Suppose W_{1max} is 12 msec and W_{2max} is 13 msec. These values can be considered the periods, or cycle times, of the periodic processes responsible for the waiting times W_1 and W_2 , respectively. Thus, these two processes will pass in and out of phase every 156 (12×13) msec. Now, consider how W_1 , W_2 , and their phase relationship are related to the stimulus sequence used in response-stimulus synchronization. The occurrence of P_1 can be considered independent of W_1 for several reasons. First, trials are subject-paced such that the intertrial interval varies greatly with respect to the cycle time responsible for W_1 . Second, there is no evidence to suggest that the conscious decision to elicit an initiation response, R_i , is dependent in any way on the central process responsible for W_1 . And third, even if one postulated a relationship between W_1 and R_i triggering, its characteristics would be lost because of two sources of temporal variability interposed between the triggering of R_i and the occurrence of P_1 . One source is due to the variable efferent delay between the central response trigger and the overt response, R_o , and the other is due to the experimentally introduced, variable foreperiod between R_i and P_1 onset. Both sources are random and relatively large compared with the cycle time of W_1 . Therefore, the occurrence of P_1 onset can be considered independent and random with respect to the W_1 cycle, resulting in a uniform distribution of W_1 waiting times.

How is W_2 independent of W_1 ? Well, since P_1 onset is registered centrally at some random point in time with respect to the W_1 cycle and since W_1 and W_2 have slightly different cycle times, knowing at what point in the W_1 cycle P_1 is registered provides no information about what part of the W_2 cycle will be intersected after the fixed deterministic delay. Consequently, W_1 and W_2 can both be considered independent and uniformly distributed. As a result, the convolution of these two distributions will generate a unit of central temporal variability that is basically triangular, that is, as long as the values of W_{1max} and W_{2max} are not too dissimilar. A slight difference in their periodicities only produces a slight bluntness in the peak of the triangular distribution. Finally, since P_1 onset, W_1 , and W_2 can all be considered independent, no autocorrelation should exist between trials, which is consistent with the data.

After information transfer to the output buffer, the output stage takes control, triggering the appropriate action (finger withdrawal), which, after an efferent delay, is manifest in the overt response, R_o . As discussed earlier, these efferent latencies are assumed to be approximately normally distributed with a relatively small variance.²⁰ In the model, the logistic distribution is substituted for the normal because of its marginal superiority in mathematically representing this kind of variability.

The ultimate goal of this sequence of processing stages is to produce a response that is perfectly synchronous with P_2 onset. However, because of variability incurred at various stages, the best performance that can be realized involves centering the response latency distribution about the time-point corresponding to P_2 onset and minimizing the variance of the various stochastic components.

The overall response latency distribution is given by the convolution of the component distributions. In the model this involves convoluting the distributions associated with W_1 , W_2 , and the efferent latency. Although $W_{1\max}$ and $W_{2\max}$ are not exactly the same, for purposes of the initial modeling they were considered identical. Thus, the convolution of the distributions of waiting times produces a triangular distribution spanning a range of $2 \times W_{\max}$. When this is further convoluted with a relatively low-variance, logistically distributed component, it produces a distribution that is still basically triangular, but with short tails and a slightly blunted peak.

The general shape of distribution generated by the model seems to characterize the data quite well, but a more rigorous test of the model's ability to represent the data was obtained by mathematically testing the goodness-of-fit between the distribution function generated by the model and the cumulative probability of response distribution derived from the data.

Parameters of the model, representing the variable components, were estimated by allowing them to vary over a calculated range while repeatedly testing for goodness of fit via the minimizing χ^2 technique. The two parameters estimated consisted of W , which represented the average of $W_{1\max}$ and $W_{2\max}$ (the maximum times required for information transfer), and b , which represents the standard deviation of the efferent response latency distribution. Values for W and b were covaried because they were constrained by the overall variance of the observed distribution. Variance of the triangular distribution is given by $W^2/6$ and variance of the logistic distribution is represented by b . Thus, the equation for the overall response latency variance V_{total} is:

$$V_{\text{total}} = W^2/6 + b^2$$

Therefore, the constraints are clear. If $b = 0$, then:

$$W_{\max} = \sqrt{6 \times V_{\text{total}}}$$

and conversely, if $W = 0$, then:

$$b_{\max} = \sqrt{V_{\text{total}}}$$

Moreover, for any W chosen in the range from 0 to $W_{\max}$ the value of b is fixed by the following equation:

$$b = \sqrt{V_{\text{total}} - W^2/6}$$

FIGURE 4 shows the excellent representation of the data provided by the model. The points on the figure are the actual data obtained from a well-practiced subject and the line is the best fit to this data provided by the model. It is evident that the model provides an excellent description of the data with a χ^2 value of 4.81 on 9 degrees of freedom. In this case, the estimates of W and b are 11.4 and 3.6 msec, respectively. This produces an estimate of overall variance of 34.7 relative to 34.6 msec² obtained from the data.

FIGURE 5 presents the results of a similar analysis based on the lowest variance data obtained in the experiments. I have added quartile lines to emphasize the amazing accuracy with which humans can perform response-stimulus synchronization. Mean response latency is only 0.3 msec longer than the P_1P_2 interval (460 msec) and the

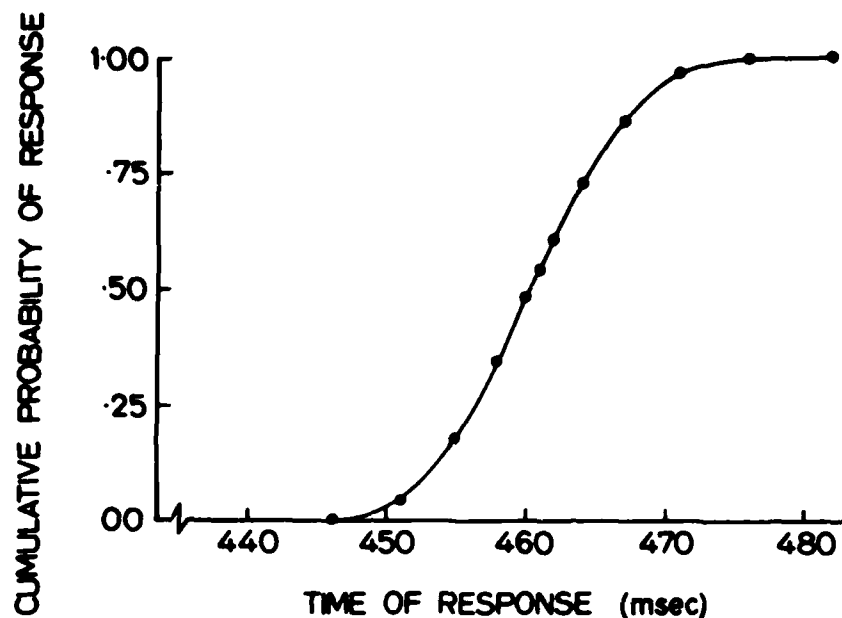


FIGURE 4. Graph showing goodness-of-fit of model to the data from sessions 266-270 for subject G.H. The *solid circles* are the data points and the *line* is the psychophysical function predicted by the model. Parameter values are given in the text. P_1P_2 interval is 460 msec.

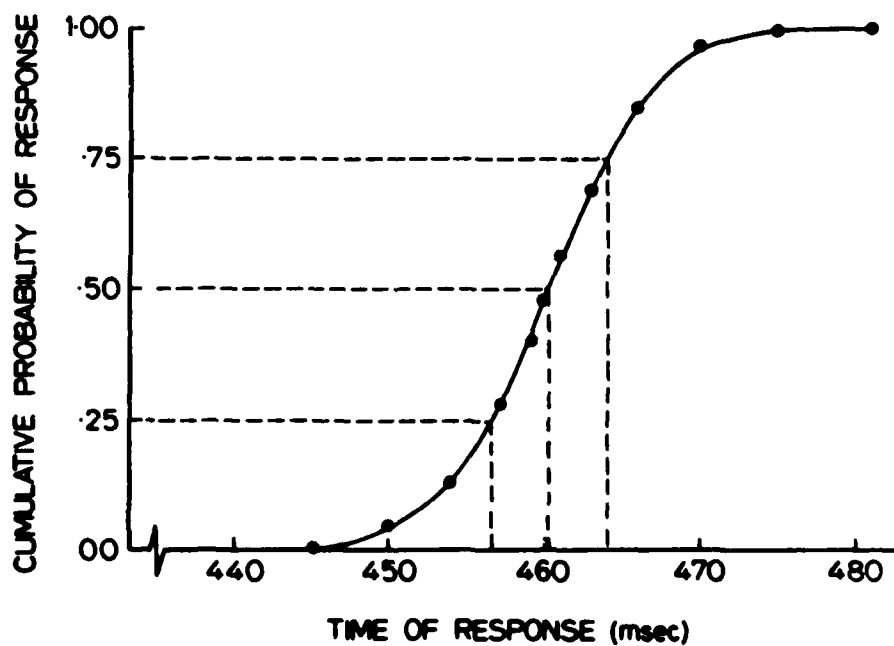


FIGURE 5. Goodness-of-fit of model to the data from five lowest-variance sessions for subject G.H. The *solid circles* show data points and the *line* the psychophysical function predicted by the model. P_1P_2 interval is 460 msec. *Dotted lines* indicate the quartile response times.

spread of the distribution of response latencies is small, with 50% of all responses falling within a 7-msec time window. Expanding the time window to 19 msec accounts for more than 90% of the responses and, in fact, all responses in this analysis fall within ± 16 msec of the mean. This distribution represents five sessions worth of data or 1500 response latencies. The estimates for W and b for this set of data are 9.1 and 4.18, respectively. These values provide an estimate of overall variance of 31.3 msec², which is exactly what was obtained from the data.

Parameters were extracted not only from stable data obtained after much practice, but also from data obtained during acquisition when the response latency variances were decreasing. Results of these analyses suggest that the value of W changes little while b exhibits substantial decreases to account for the reduced variances.

In spite of the fact that the model fits the data extremely well, the method of extracting parameters seemed to lack power because there was usually a range of W and b values that produced near optimal fits to the data. For example, in one analysis b values ranging from 3.5 to 5.0, in combination with their corresponding values for W , all provided reasonably acceptable representations of the data. As a result, I was concerned about the ability of the minimizing χ^2 procedure to extract meaningful estimates of the underlying parameters involved.

To further exploit the model, I decided to simulate subjects' response latency distributions, using parameters extracted from the original data, by randomly generating response latencies which summed the delays associated with each of the model's variance components. This simulated data was then analyzed to try and recover the parameters used in generating it. This would provide some indication of the power of the parameter-extraction procedure employed.

Results of these simulations were quite rewarding. Parameter estimates based on the simulated data never exceeded 10% error from those parameters used to generate the data, even with the apparent lack of power associated with the parameter-extraction technique, and typically the correspondence was almost perfect. Two examples are shown in FIGURE 6. The left panel shows a simulated response latency distribution generated with $W = 11.4$ and $b = 3.6$. The corresponding parameters extracted from this data were 11.6 and 3.5, respectively. The χ^2 statistic for the fit is 4.7 on 11 degrees of freedom. The second example, shown in the right panel, was based on a set of data generated with $W = 11.6$ and $b = 3.5$. In this case, the recovered value for W was identical at 11.6 and the b value was very close at 3.35, with a χ^2 value of 5.27 on 10 degrees of freedom. The overall variances associated with these two distributions are 35.1 msec² for the left-hand distribution and 33.7 for the other.

All the analyses to this point have been using an average value for W_1 and W_2 . However, as pointed out earlier, it is imperative to the assumption of independence between these two components that they be slightly different. To test whether this restriction would affect theorizing, several sets of data were simulated using different values for W_1 and W_2 . As also pointed out earlier, it is only necessary for W_1 and W_2 to differ by a millisecond or less for the argument of independence to hold. Consequently, data sets were generated in which W_1 and W_2 differed by 1 msec. The results were almost identical to those obtained when the average values of W_1 and W_2 was used. In fact, FIGURE 6 could also be used to present these data because the differences in the distributions and parameters extracted are less than 1 percent.

Thus, the method used for parameter extraction does not appear to impose any serious restrictions on our ability to estimate W and b . Moreover, it is interesting to note that the estimates of b that were calculated, where b^2 represents the efferent delay variance, agree well with estimates provided by Wing and Kristofferson,⁹ and the estimate of W is in the same range as the minimum unit of quantal time that Kristofferson has proposed.³ However, this latter correspondence does not imply a

single "clock." In fact, the mechanism proposed to produce independence assumes two different periodicities.

One consistent finding was that efferent latency variance, estimated by b , tends to contribute proportionately less to the total variance as a function of practice. In terms of the response latency distribution, it means that the shape should become more triangular as practice continues and, to some extent, this can be seen in the data. The coefficients of kurtosis are generally around 3.2 at the beginning of practice and decrease over time to values under 3.0. The coefficient of kurtosis for the logistic distribution is 4.2 and for the triangle is 2.4. Therefore, the trend towards a reduction in the coefficient of kurtosis, especially in the range observed, is consistent with the notion of a transition in shape from logistic to triangular. However, this information should only be taken as corroborating evidence because the change in coefficient values was rather small and inconsistent in a few instances.

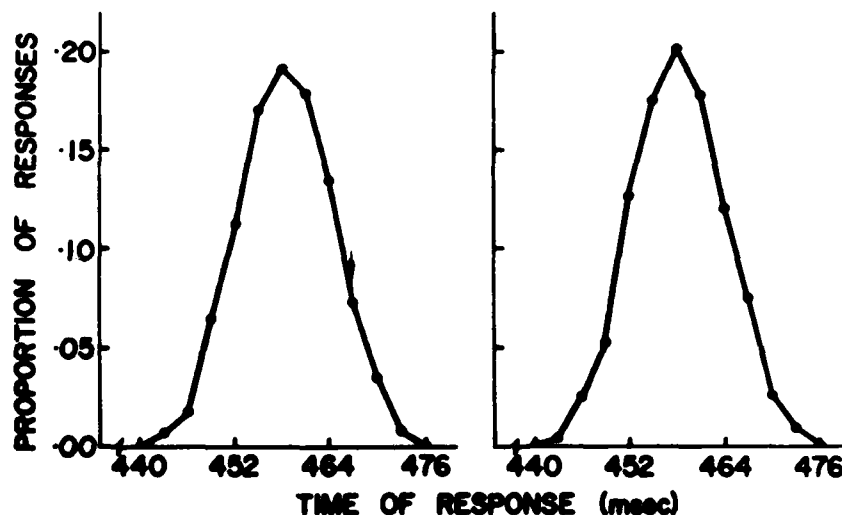


FIGURE 6. Relative frequency distributions for two sets of simulated data. Left panel shows distribution generated from the model (see text), with $W = 11.4$ and $b = 3.6$. Extracted parameter values based on these data were $W = 11.6$ and $b = 3.5$. Right panel shows similar distribution generated with $W = 11.6$ and $b = 3.5$. Recovered parameter values were $W = 11.6$ and $b = 3.4$. χ^2 for goodness-of-fit of model to the left and right distributions was 4.7 on 11 d.f. and 5.3 on 10 d.f., respectively.

Up to this point I have been referring to the parameter b as representative of efferent variability. But the analyses indicate that most of the improvement with practice is reflected in reductions in b . Because several thousand trials are required to attain asymptotic performance, it is difficult to understand how a simple finger withdrawal response could continue to be refined after such a number of trials. Consequently, I would like to introduce a further extension to the model that addresses this problem as well as that of violating the assumption of steady-state behavior, as mentioned earlier.

In reality, the state of the central mechanism governing response trigger timing probably changes slowly over time, being affected by concurrent cognitive activity and changes in the physiological state of the organism. Thus, internal conditions could be viewed as if they were in a state of continual flux. This would necessitate a dynamic

process for fine-tuning, or updating, the central timing stage based on feedback about recent successes and/or failures in synchronization.

The result of this continual updating of the system introduces a new source of variance that was not previously considered. As the internal state of the organism changes, timing will be affected such that the response triggers will begin to occur too early or too late on the average. This information, provided by the available feedback, allows the central stage to alter the timing process accordingly in order to maintain accurate synchronization.

Obviously the accuracy of such a feedback loop is determined, in part, by the amount of information considered in determining the extent of alterations to be made. To be most accurate the subject should try to integrate, or average, as much information from preceding trials as possible. Such a notion is supported by the sequential dependency analyses that were applied to the experimental data. Those analyses indicate that feedback information becomes more reliable with practice, suggesting the experienced subject integrates increasingly more information into the updating decision process. Further increases in updating accuracy probably also stem from the subject's learning to interpret the feedback information better.

If this new variance component is considered to be logistically distributed, then it would be combined in our estimate of b . Thus, b may actually encompass two sources of variation that cannot really be separated because they are similarly distributed. Probably the most reasonable explanation of decreases in the value of b with practice would share the responsibility between decreases in efferent latency variance and increases in accuracy of updating the central components to maintain synchronization with the "real world." For example, the large decreases in b that are observed early in practice could be largely a function of response learning, while after some initial training, continued decreases in b could reflect the improved updating abilities of subjects. However, when asymptotic performance is attained, indications are that b provides a good estimate of minimum efferent latency variance because the updating process has been so refined as to approximate a steady-state system.

What happens without the feedback? Well, the data show that performance deteriorates rapidly and strong sequential dependencies begin to emerge. In terms of the model, lack of feedback prevents the updating process from operating, which results in mean response latency varying over time as a function of changes in the internal state of the organism. This produces a slow wandering of the mean, thereby inflating overall response latency variance and introducing autocorrelations between successive responses. Decrements in performance are probably further enhanced by degradation of the memorial representation of the P_1P_2 interval. Without feedback, there is no way to refresh the memory for the synchronization interval.

If feedback helps to synchronize the central timekeeping process with the passage of physical time, then one would expect individual differences to be minimized when feedback is available. This is exactly what was observed in these experiments. When feedback was present, all subjects, independent of sex and P_1P_2 interval, exhibited very similar performance. However, in the absence of feedback, large individual differences developed in terms of mean errors of synchrony, response latency variances, and autocorrelations.

SUMMARY

To summarize this discussion, I have presented several pieces of evidence in support of the proposed model. Substantial support comes from the fact that when parameter estimates obtained from independent areas of research are substituted into the model.

it predicts asymptotic variance levels almost identical to those observed. The model also provides a remarkable mathematical representation of the data and can account for independence of mean and variance over the range of P_1P_2 intervals tested. Extensions made to the model provide a locus for the central variance components in the S-R chain and also provide a mechanism to produce independence between the central components by assuming two different periodicities. Finally, not only did the model provide a good mathematical representation of the data, but also the parameters were theoretically appealing for further postulating about central mechanisms and processes underlying response-stimulus synchronization. This last point is important because we also tested several other models, which I have not discussed, that provided acceptable mathematical representations of the data, but none had parameters that were theoretically meaningful.

As a final point, I would like to reexamine the traditional distinction made between subjective and objective time. Subjective time has often been thought of as a dimension of experience only, in which the nature of the activities occurring during a period is the major determinant of the phenomenal duration, rather than movements of hands on a clock.²¹ In fact many investigators, back to the time of William James,²² have been intrigued by these alterations in perceived duration produced by varying the physical events generating our subjective experiences. Obviously these are important aspects of cognitive functioning to understand, but it is also intriguing to find out that under some circumstances there is no transformation made between physical and psychological time. The two are the same. The characteristics of these circumstances, however, have just recently begun to emerge from synchronization and duration discrimination studies. Consequently, it is important that this mode of information processing receive further investigation because possibly it provides the crucial link between our minds and our environment in "real" time.

REFERENCES

1. KRISTOFFERSON, A. B. 1976. Low variance stimulus-response latencies: Deterministic internal delays? *Percept. Psychophys.* **20**: 89-100.
2. KRISTOFFERSON, A. B. 1977. A real-time criterion theory of duration discrimination. *Percept. Psychophys.* **21**: 105-117.
3. KRISTOFFERSON, A. B. 1980. A quantal step function in duration discrimination. *Percept. Psychophys.* **27**: 300-306.
4. MICHON, J. A. 1967. *Timing in Temporal Tracking*. Institute for Perception (RVO-TNO). Soesterburg, The Netherlands.
5. KRISTOFFERSON, A. B. 1967. Attention and psychophysical time. *Acta Psychol.* **27**: 93-100.
6. CREELMAN, C. D. 1962. Human discrimination of auditory duration. *J. Acoust. Soc. Am.* **34**: 582-593.
7. TRIESMAN, M. 1963. Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychol. Monogr.* Vol. **77** (13, whole no. 576).
8. GETTY, D. J. 1975. Discrimination of short temporal intervals: A comparison of two models. *Percept. Psychophys.* **18**: 1-8.
9. WING, A. M. & A. B. KRISTOFFERSON. 1973. Response delays and the timing of discrete motor responses. *Percept. Psychophys.* **14**: 5-12.
10. ALLAN, L. G. 1975. The relationship between the perception of successiveness and the perception of order. *Percept. Psychophys.* **18**: 29-36.
11. WING, A. M. 1974. The timing of interresponse intervals by human subjects. *Diss. Abstr. Int.* **35**: 4237B.
12. BOTWINICK, J. & L. W. THOMPSON. 1966. Premotor and motor components of reaction time. *J. Exp. Psychol.* **71**: 9-15.

13. SASLOW, C. A. 1972. Behavioral definition of minimum reaction time in monkeys. *J. Exp. Anal. Behav.* **18**: 87-106.
14. MICHAELS, A. A. 1977. Timing in Sensorimotor Synchronization. Unpublished Master's thesis, McMaster University, Ontario, Canada.
15. TYLDESLEY, D. A. & H. T. A. WHITING. 1975. Operational timing. *J. Hum. Movement Stud.* **1**: 172-177.
16. ADAMS, J. A. & L. R. CREAMER. 1962. Proprioception variables as determiners of anticipatory timing behavior. *Hum. Factors* **4**: 217-222.
17. JONES, B. 1972. Outflow and inflow in movement duplication. *Percept. Psychophys.* **12**: 95-96.
18. KEELE, S. W. 1973. Attention and Human Performance. Goodyear Publishing Company. Pacific Palisades, CA.
19. SCHMIDT, R. A. & R. W. CHRISTINA. 1969. Proprioception as a mediator in the timing of motor responses. *J. Exp. Psychol.* **81**: 303-307.
20. MEIJERS, L. M. M. & E. G. J. EIJKMAN. 1974. The motor system in simple reaction times experiments. *Acta Psychol.* **38**: 367-377.
21. ORNSTEIN, R. E. 1969. *On the Experience of Time*. Penguin Books. Harmondsworth, England.
22. JAMES, W. 1908. *Principles of Psychology*, Vol. 1. Holt. New York, NY.

The Perception of Temporal Events^a

D. ALAN STUBBS, L. R. DREYFUS,
AND J. G. FETTERMAN

*Department of Psychology
University of Maine
Orono, Maine 04469*

Much of the research on animals' temporal discrimination has an implied framework that treats the phenomena from a sensory process point of view.

Recent years have seen the development of a variety of psychophysical techniques.^{1,2} For example, animals may be presented with different durations of a stimulus, with one choice response reinforced if the duration had been short and a second choice response reinforced if the duration was long. In addition, there has been an increase in research using reinforcement schedules that is psychophysical in orientation; orderly relations between schedule value and behavior have been observed under a wide variety of temporally defined schedules.^{3,4} This line of research has provided valuable information about Weber's law, logarithmic and power relations, and subjective time scales. While this sort of research has produced basic and necessary information, most of it has treated stimulus duration as a simple dimension that is similar to sensory dimensions such as frequency or intensity. In most of the research, the durations to be discriminated are composed of a single unchanging light or tone. The stimulus component has been de-emphasized so that orderly relations could be obtained and so that the operation of a sensory-based timing process could be assessed.

Not only the procedures, but also current theories, imply a sensory framework. Recent theories are concerned with the timing process and, in general, all attempt to separate the timing process and factors affecting it from other processes and aspects of the situation.^{3,5,6} The different theories have different emphases. Gibbon's scalar expectancy theory, for example, emphasizes the orderly psychophysical relations that transcend individual procedures in an attempt to delineate the temporal scaling process that is common to the different situations. The internal clock theory of Church and his students has emphasized the "clock" and has focused on factors that influence the way in which the clock functions. Although no one argues that the timing process is simple, the treatment of timing does indicate a sensory process framework. The theories are typically described in cognitive terms and the theories take an information-processing approach, but the roots seem to come from a sensory-based orientation. In particular, the internal clock theory seems to treat the clock as a receptor that registers the passage of time. With sensory dimensions such as brightness there are correlated physiological processes; the clock seems to serve a similar though hypothetical function as a time sense.

An alternative orientation places temporal discrimination within the framework of perception. This framework places animals' time perception within the context of other complex perceptual phenomena such as space perception, motion perception, and form perception. A perceptual framework suggests analogies between animal time perception and other perceptual processes and indicates that the study of animal timing might

^aAuthorship of this paper is equal. The research was funded in part by a Faculty Research Fund Award from the University of Maine to D. Alan Stubbs.

well profit from the lessons learned in these other areas. Stimulus durations may be arranged easily (more easily than most stimulus continua), orderly psychophysical findings might occur when durations are used, and inferences about a timing process can be made. But the ease of arranging the stimulus dimension and the orderly relations may be misleading and deceive us into viewing the problem in a particular way. With space, movement, and form perception, continua have also been established, elegant psychophysical experiments have been conducted, and valuable results have been obtained, but more is involved in all of these areas. With space perception there are the many "cues," both stimulus and organismic; prior experience tells us about the distance of things; different information is provided to stationary and moving observers; and there are higher-order invariants and relations when observers move about in a world of textured surfaces. In light of all of these factors, the space perception literature does not have an analogue to timing; that is, researchers do not talk about a "spacing" process. Research such as that with the Howard-Dolman apparatus (in which people judge the relative distance of two rods against a homogeneous background) demonstrates the role of accommodation, convergence, and disparity. These organismic or internal cues provide the closest analogy to timing. However, the space perception literature has shown that these cues tell only a part of the story; much more is involved when humans and nonhumans move about in less restricted visual environments.

The adopting of a perceptual framework leads to the conclusion that the implied sensory framework that guides most current thinking may be limiting the questions that are asked and the research that is done. A perceptual approach argues for a change in emphasis and for increased attention in several areas. First, greater emphasis should be given to the stimuli constituting the durations to be discriminated. Most researchers recognize that temporal discrimination tasks do not measure discrimination of time, but instead the discrimination of the duration of stimuli. It would seem proper that a greater emphasis be given to the stimulus side of stimulus duration. While human research has employed a wide variety of stimulus situations, most animal research has used durations marked by lights and tones that do not change. The function of these stimuli seems to be simply that of a time marker; the stimulus does not appear to be given much importance, but rather serves as a way of bounding abstract time or setting a clock in operation. Second, most of the research has attempted to eliminate environmental events in duration tasks. With people, time judgments are influenced by watches, the sun, the amount of work done, the variety and structure of events, and the like. But these conditions are often eliminated in research, particularly that with animals; they are eliminated in part because they are viewed as contaminating features that might provide extraneous cues to interfere with "pure" timing. The perception of time is more properly considered as the perception of temporal events and as such resembles in some ways the perception of movement and perception of change.^{7,8} Unfortunately, different research strategies separate the areas: with movement and change, researchers are searching for the relevant stimuli; with time, researchers (particularly animal researchers) typically hide or limit the relevant temporal events. Use of unchanging stimuli in duration tasks reduces temporal structure in the environment so that the only events remaining are internal, which inevitably leads to a view of these events as crucial. The situation may be like that with space perception in the past, when accommodation, convergence, and disparity assumed a role of importance when experiments eliminated most of the features of a textured world. Third, the perceptual framework suggests that a wider range of tasks, including tasks that are more complex, be used. Although a variety of procedures are currently available, these procedures are primarily variations on the basic human psychophysical techniques. These sensory-like techniques contribute to the sensory

process framework for considering research findings. What is needed is a more varied set of tasks to deal with questions like those raised in human time perception and like those raised in other areas of perception.

The two experiments to be reported have their origins in our consideration of temporal discrimination within the perceptual framework.

EXPERIMENT 1: THE AUDITORY-VISUAL DIFFERENCE IN PIGEONS

The purpose of the first experiment was to see whether different stimuli would influence performance in a duration task. A robust finding in the area of human time judgment is that people judge sounds as longer than lights of equal duration.⁹ This finding holds across different procedures and variations in method and thus seemed to be a good starting point in the investigation of stimulus factors as they affect temporal judgments of animals.

The procedure used for the task was based on the free-operant scaling procedure of Stubbs.¹⁰ The scaling procedure was selected since it permits a greater flexibility of response than do discrimination procedures and thus seemed to offer a greater likelihood of differences in response when lights and sounds were presented.

Three pigeons were trained in a standard operant chamber, with daily sessions lasting until a pigeon had received 50 reinforcers. FIGURE 1 gives the outline of the procedure. Intertrial intervals separated the duration periods. The intertrial intervals were variable in length, with an average of 15 sec. In addition, a reset contingency was used such that a peck to either key would prolong the intertrial interval by 5 sec. During the intertrial interval, the houselight was on and the right key was lit by white light. After the intertrial interval, there were two types of trials, with each occurring equally. On light trials, the onset of the trial was signalled by the onset of a red light behind the left key (the houselight and white light remained on). When the red light was on, red-key responses were sometimes reinforced between 5 and 6 sec after light onset. If food was produced, an intertrial interval followed and then a new trial began. The pigeon could also peck the white key, with the result that the stimulus conditions would change with green replacing red on the left key. Once the green light was on, pecks to white had no further consequences (that is, the pigeon could not change back

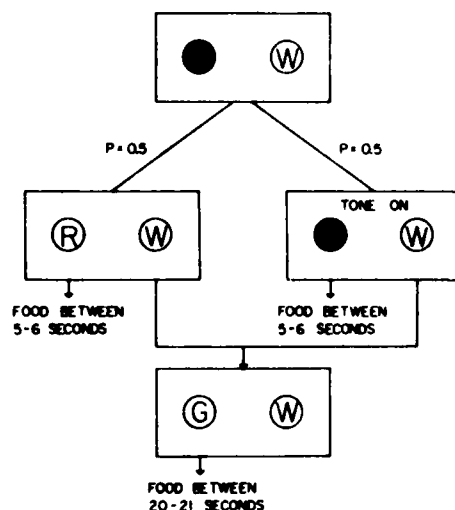


FIGURE 1. Outline of the procedure for Experiment 1. The *top* portion represents the intertrial interval, which was variable in duration. A trial began with either the onset of a red light or a tone. Food was produced intermittently for left-key responses 5-6 sec after stimulus onset. A peck to the white key changed the stimuli and, in the presence of green, left-key responses produced food intermittently 20-21 sec after trial onset. Trials ended with food or if 24 sec elapsed without food, and then a new intertrial interval began. A houselight was on at all times except during food delivery.

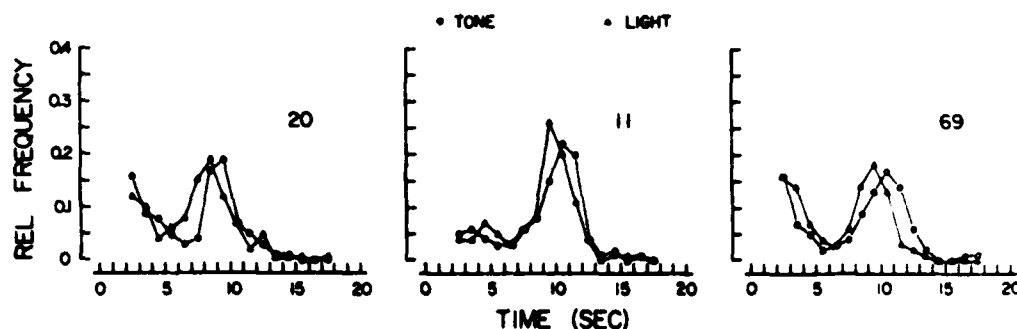


FIGURE 2. Relative frequency of a changeover as a function of time since a stimulus period began. Data were computed separately and are shown for both light and tone trials. Food was delivered between 5 and 6 sec in the presence of red light or the tone and between 20 and 21 sec in the presence of green. Data are based on the total number of changeovers in the last five sessions.

to red). Pecks to the green key were sometimes reinforced between 20 and 21 sec of trial onset. If food was not produced, the trial ended after 24 sec. The contingencies for tone trials were the same except that a tone (produced by a Sonalert) was presented in place of the red light. Trials began with the onset of the tone, which was the only change in stimuli after the intertrial interval. The houselight and white key light were kept on during intertrial intervals so that changes in visual stimuli would not accompany the onset of the relevant auditory stimulus. On tone trials, a peck to the dark, "tone" key sometimes produced food between 5 and 6 sec after tone onset. A peck to the white key turned on the green key light and turned off the tone. The contingencies in the presence of green were the same as on light trials.

Food was made available on some trials only. Once available, food was assigned equally for short and long responses to ensure an equal distribution of reinforcers. For reinforcers assigned for short responses, half were assigned during tone trials and half during light trials. If a reinforcer was assigned but missed, the reinforcer was assigned again on the next appropriate trial. This reinforcement procedure fixed relative reinforcement frequency under the different conditions to make sure that any performance differences would not be due to differences in obtained number of reinforcers.

Basically the procedure provided reinforcement at a short and a long time after stimulus onset. The major difference in trials was whether a light or a tone was associated with the early portion of the time period, while a light was always associated with the later portion. The question was whether the light-tone difference early in the time period would affect the time at which the animals would change to green.

FIGURE 2 shows the relative frequency of a changeover from red to green and from tone to green in different time classes. The general finding, in agreement with previous research, is that the frequency of changing over was highest at times that approximated the geometric mean (10–11 sec) of the two intervals at which reinforcement was provided (5–6 sec [short] and 20–21 sec [long]). The forms of the distributions were similar on light and tone trials, except that the distributions are shifted to the right under tone conditions. The difference in the distributions suggests that the time to food in the presence of the tone was judged as longer than that in the presence of the red light. Although red-key and "tone"-key responses both produced food at the same interval after stimulus onset, perception of time to food in the presence of the tone as longer would result in a longer time spent before changing to green.

The results agree with the findings on human time judgments and suggest that the auditory-visual difference holds for pigeons as well as people. While the results suggest an auditory-visual difference, we must be cautious in the interpretation. No attempt was made to equate the intensity of the light and sound. Although the auditory-visual difference holds for people even with changes in stimulus intensity,¹¹ our results may have depended on differences in intensity. Also, light and tone trials differed with regard to whether the pigeons pecked on a red or dark key (although response rates were similar in both cases). Probably these factors are not responsible for the difference in performance, but the matter needs future corroborative research.

Whatever the reason for the difference, the results demonstrate that the stimuli that demarcate temporal events affect performance. The results join a small body of literature showing that stimuli can have varied effects on temporal judgments. Mantanus,¹² for example, trained pigeons on a duration-discrimination task and found that accuracy was higher when durations were "filled" than when they were "unfilled." Similarly, Spetch and Wilkie¹³ observed higher accuracy when durations of food access were used as opposed to durations of a light. Stulz *et al.*¹⁴ found that discrimination performance was affected not just by the stimuli that make up durations but also by the stimuli that bound durations. They trained pigeons on a schedule in which fixed-interval performance resulted sometimes in food and sometimes in a briefly presented stimulus. Occasionally the schedule was interrupted by a choice situation: one response was reinforced if only a short portion of the interval had elapsed, while a second response was reinforced if a longer portion had elapsed. The main result was that accuracy of choice was lower if the interval was preceded by the brief stimulus rather than food. The few available results suggest that the stimuli that make up temporal events have varied effects that include changes in accuracy and changes that suggest differences in perception of different temporal events. Only detailed attention to a whole range of stimulus conditions will reveal the range and extent of stimulus factors.

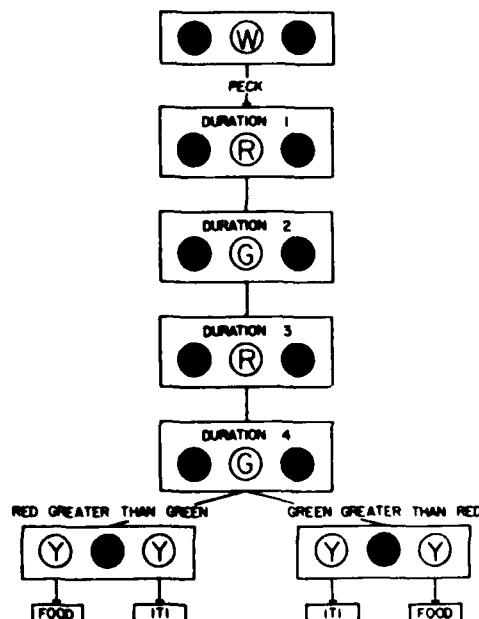
EXPERIMENT 2: TEMPORAL INTEGRATION

There are numerous procedures for the study of animals' temporal discrimination, but in most the consequences depend on the passage of a single interval. Fetterman and Dreyfus developed an alternative procedure in which performance depended on more than the passage of a single interval of time.¹⁵ The procedure involved the comparison of two durations with different choice responses reinforced depending on which duration was longer. Pigeons were presented with a red light of one duration followed by a green light of a second duration; then two choice-key lights were turned on and a response to one key was reinforced if the duration of red had been longer than green, while a response to the second was reinforced if the duration of green had been longer than red. The pigeons were accurate in discriminating many different durations, with accuracy a function of the relative difference in the durations.

The present experiment is derived from the method of Fetterman and Dreyfus and represents an increase in the complexity of the task. Pigeons were given four rather than two durations, in the order red-green-red-green, with the duration of each color changing from trial to trial. As before, the pigeons' task was to give one response if the duration of red was longer than green and another if the duration of green was longer than red; the added feature was that the pigeons were required to integrate or add together the two red durations and the two green ones and to respond on the basis of whether the total red duration was shorter or longer than that of the total green.

FIGURE 3 outlines the procedure. For each trial, the center key was white initially and the side keys dark. One response on the center key started the series of stimulus durations on the center key with each changing independently of behavior. Each series consisted of four durations: red, green, red, green. Each duration was arranged by a probability gate set at a value of 0.10, which was pulsed once every 0.5 sec. The average time of each duration was thus 5 sec, with a range between 0.5 and 16 sec. This system provided for more than a million different combinations of durations and allowed for the total of the four durations to be as short as 2 sec or as long as 64 sec. After completion of the fourth duration of the series, the center-key stimulus went off and the side-key stimuli came on. If the total of the two red durations was longer than that of green, a left-key response was correct; if the total of green was longer than red, a right-key response was correct. Correct responses either resulted in a 10-sec intertrial interval (during which all lights were off) or intermittently produced food. Food was assigned randomly from reinforcer to reinforcer either for a correct left-key or for a

FIGURE 3. Outline of the procedure for Experiment 2. A response to the white key started a series of four durations, with the series changing from trial to trial. At the completion of the duration series, two side-key stimuli were lit. One response was reinforced if the total of two red durations had been longer than the total of green; the other response was reinforced if green had been longer than red. Incorrect responses produced an intertrial interval during which all lights were off.



correct right-key response. If a reinforcer was assigned to the left (right) only a correct response to that key could produce food; when a reinforcer was assigned to the left (right), correct responses to the right (left) key simply produced the intertrial interval. This procedure was instituted to hold the number of reinforcers equal for left-key and right-key responses. The net result of this procedure is that roughly half of the correct responses resulted in food. All incorrect responses produced the 10-sec blackout. In addition, in those cases in which the total red and green durations were equal, choices would be neither correct nor incorrect, so choices always produced a blackout under these circumstances. Each session consisted of a series of trials that lasted until a pigeon received 75 reinforcers.

FIGURE 4 shows the probability of a right-key response (reporting green as longer than red) as a function of the relative difference in the durations of red and green and shows orderly functions relating probability of a right-key response to changes in the relative durations of red and green. The functions are not symmetrical around the one

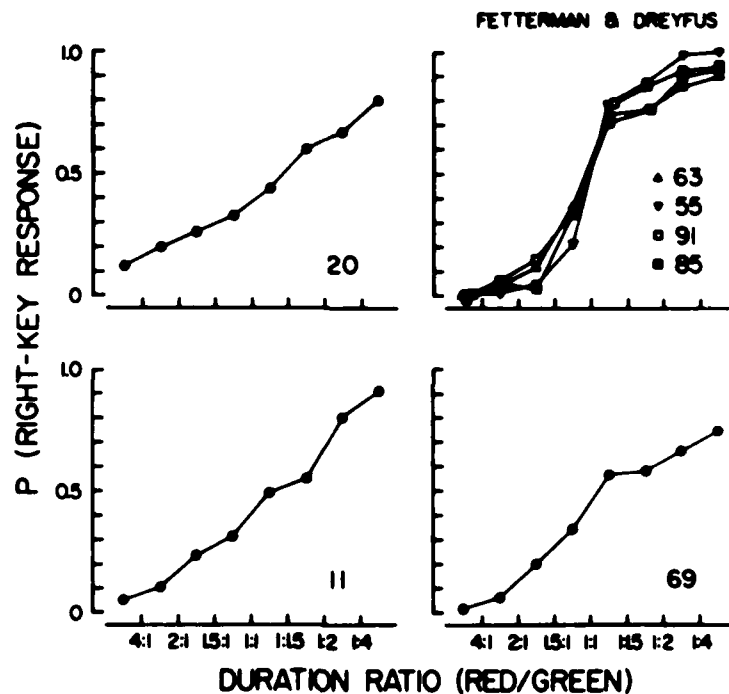


FIGURE 4. Probability of a right-key response ("green longer than red") as a function of the duration ratio of red to green. The first point shows performance when the duration of red was greater than four times that of green; the second shows performance when red was two to four times longer than green, and so forth. The data comprise many different duration series in which data from the last five sessions were totaled; data included approximately 1000 problems per pigeon. The upper right panel shows equivalent data from an experiment by Fetterman and Dreyfus in which performance depended on two rather than four durations.

to one ratio; there was a bias to respond left or, stated differently, accuracy was higher in cases when red was longer than green as compared with similar cases when green was longer than red. Calculations on the few instances (5-6 per person) in which red and green durations were equal show that the pigeons responded more often on the left key in these cases. The pattern of performance persisted in spite of the procedural arrangement of holding the relative number of reinforcers equal for the two choice responses.

The upper right hand panel of FIGURE 4 gives the comparable data from the experiment by Fetterman and Dreyfus, who used a similar procedure except with only two durations for comparison rather than four. Comparison of performance in the two experiments reveals steeper ogival functions in their experiment, indicating that performance was more accurate. This difference is not surprising considering the greater difficulty of the present experiment.

FIGURE 4 included performance across a wide range of combinations of durations. A reasonable question is whether performance was similar when the durations were short and when they were long. Notions of memory might suggest that accuracy would be lower when the durations were long since a greater time would elapse between the early durations in the series and a choice. FIGURE 5 categorizes the data into two classes of duration that depend on the total time of the duration series, with the classes being divided in terms of the average combined time of the four stimuli.

FIGURE 5 indicates that the two sets of data were comparable and that accuracy

was similar whether the durations were short or long. The two functions are similar for pigeons 20 and 69, with a great deal of overlap. For pigeon 11, the open circles lie above the solid circles in most comparisons. The difference in performance appears to be one of bias rather than sensitivity. Pigeon 11 was relatively more likely to respond left when the total duration was long and right when it was short, but accuracy was comparable in both cases.

Another question concerned the relative contribution of the first and second red-green duration pairs. The reinforcers were arranged with respect to the total red and total green durations, but the question can be raised whether choice responses were controlled to a greater degree by one red-green pair than the other. Consider the second red-green pair. Sometimes the red-green relation was the same in both red-green pairs (for example, red longer than green in both pairs) and sometimes the relation was opposite in the two pairs. In addition, the duration of the second pair was sometimes longer than that of the first pair and sometimes shorter. If the second pair was longer, the contribution of this pair to the total red-green duration would be greater than the first; as a result, the second pair, although opposite to the first, would provide the greater contribution to the total red-green ratio and hence the second red-green relation would probably correspond to that of the total. The point is that a pigeon, responding on the basis of only the second red-green pair of a duration series, could be correct more often than not due to the interrelations between the different durations.

FIGURE 6 provides the relevant data by showing probability data when the red-green relation of the first red-green pair (top portions) or of the second pair (bottom) was in the same or the opposite direction as the relation of the total red and green. In the top portions, for example, the solid circles show performance when red was longer than green both in the first pair and total relation, and when green was longer in both cases; the open circles show performance when the first red was longer than green while the total green was longer and when the first green was longer than red while the total red was longer.

The most general finding is that performance changed as a function of the total red-green duration ratio regardless of whether a similar or opposite relation occurred in the first or the second red-green pair. The data indicate that performance was not controlled exclusively by either the first or second red-green pair. At a more specific

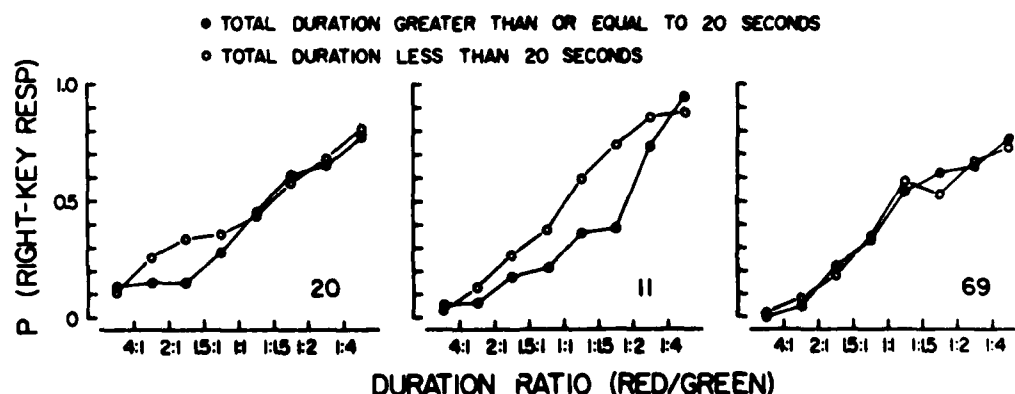


FIGURE 5. Probability of a right-key response as a function of the duration ratio of red to green. Performance is shown separately for problems in which the total time of the four durations was 20 sec or longer (*solid circles*) or less than 20 sec (*open circles*). The data are based on totals of the last five sessions, with roughly half of the problems 20 sec or longer and half shorter than 20 sec.

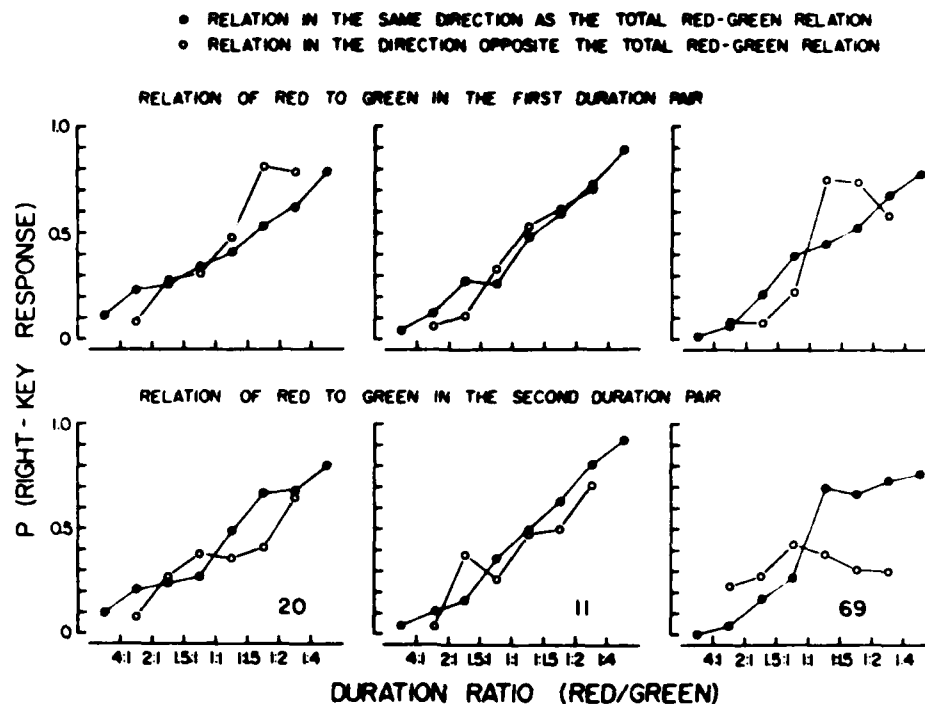


FIGURE 6. Probability of a right-key response as a function of the duration ratio of red to green. Presentation of the data emphasizes the first red-green duration pair (top) and the second (bottom) regarding whether the relation in that pair was in the same or opposite direction as the total red-green relation. In the top panels solid circles represent performance when the duration of red (green) was longer than green (red) both in the first pair and also in the total relation. The open circles represent performance when duration of red (green) was longer than green (red) in the first pair, but with green (red) longer in the total relation. The bottom panels show similar relations computed with respect to the second red-green pair. Open circles are not shown at extreme ratios since few instances (generally less than ten in the five sessions) occurred.

level, the data indicate a somewhat greater contribution by the second red-green pair. The data of pigeon 69 are instructive since this bird showed greatest control by the second red-green pair. The bottom panel for this pigeon indicates an increasing function when the second pair relation was in the same direction as the total, but a much flatter function when the two relations were opposite in direction. The small change in performance and generally lower accuracy when the red-green relation in the second pair differed from that of the total indicate greater control by the second red-green pair. Comparison of the data for pigeons 20 and 11 shows that the second-pair functions were more comparable, with overlap in the two sets of points. However, accuracy measures derived from the comparable points indicate that even for these pigeons accuracy was higher when the second-pair relation was in the same direction as the total (68% versus 62% for pigeon 20 and 72% versus 69% for pigeon 11). The data indicate a relatively greater weighting of the second red-green pair in determining choice that ranged from slight (pigeon 11) to moderate (pigeon 20) to strong (pigeon 69). For the first red-green pair (shown in the top panels of FIGURE 6), the functions are roughly similar. If there is any difference, the difference is one of higher accuracy when the relation in the first red-green pair was opposite that of the total. The reason for this result has to do with the interrelation of the first and second red-green pairs as they affected the total. If, for example, red was longer than green in

the first pair, then green would have to be even longer in the second pair for the total green duration to be longer than red. The greater difference in the second pair (whether due to a more extreme red-green ratio or to longer durations in the second pair) would enhance accuracy, especially in light of the greater weighting given to the second red-green pair. The data both from the top and bottom portions of FIGURE 6 indicate somewhat greater control by the second red-green pair.

One bird was exposed to a condition in which the durations were increased four-fold. Only one bird was used because of the increased session length of approximately 8 hours. FIGURE 7 shows results like those in the previous figures. The left panel shows performance for all durations and also data from the prior condition for comparison. Performance was similar for the two conditions. The middle panel shows that performance was comparable whether the total duration was more or less than 80 sec. The right panel shows performance computed with respect to the second red-green pair and whether the red-green relation was the same or opposite the total red-green relation; performance was similar to that in FIGURE 6. Basically the results corroborate those of the previous condition while extending them to a longer set of durations.

The results demonstrate that pigeons can integrate and discriminate the duration of events in the task, that accuracy is not markedly affected by the total time of the four durations, and that performance is based on the combination of more than two of the four durations. The results are in general agreement with previous results in which responses depended on the duration of a single stimulus or on the comparison of two durations.^{15,16} The different results agree that discrimination accuracy is a function of relative rather than absolute differences in durations and that performance is comparable across different ranges of durations. The most obvious difference in the results is lower accuracy in the present task (for example, the comparison in FIGURE 4). That accuracy was lower is not surprising since the task was more difficult, involving the integration and comparison of several durations.

Performance was controlled by all four durations, but the evidence suggests that there was a greater weight given to the earlier durations and, at the same time, a

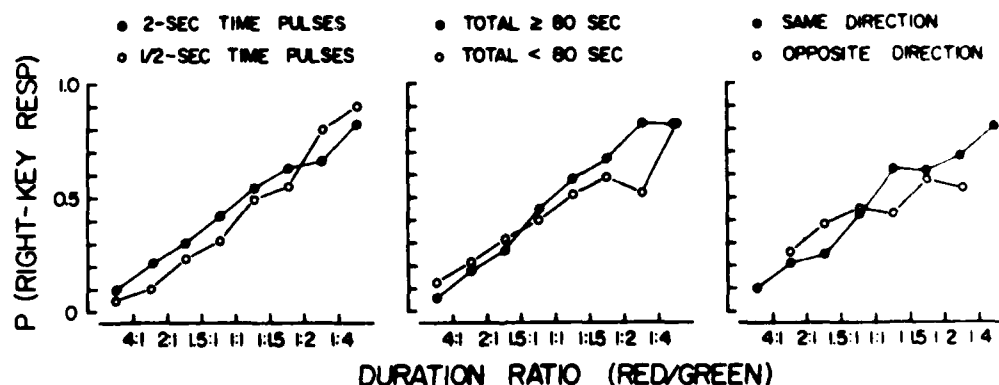


FIGURE 7. Probability of a right-key response ("green longer than red") as a function of the duration ratio of red to green. The **left panel** shows performance on all problems when the time base controlling duration was 2 sec (*solid circles*) and from FIGURE 4 when it was 0.5 sec (*open circles*). The **center panel** shows performance for problems with a total duration of 80 sec or greater (*solid circles*) or less than 80 sec (*open circles*). The **right panel** shows performance computed with respect to the second red-green pair when the relation in that pair was in the same direction as the total (*solid circles*) or the opposite direction (*open circles*). Data points are based on totals of the last five sessions.

greater weight given to the second red-green pair. Evidence that early durations were weighted more heavily comes from findings in FIGURE 4 that accuracy was higher in cases in which red was longer than green as opposed to comparable cases in which green was longer than red (for example, 2:1 versus 1:2). Also, the birds responded as if red was longer in most cases when red and green durations were equal. These results could be due to a left-key bias that was independent of the task, but this possibility seems rather unlikely since performance was similar in all three birds and since the reinforcement contingencies, which fixed relative reinforcement rate, would minimize such a bias. It is more likely that something about the task produced the difference in performance. The results suggest that the early durations might be perceived as being relatively longer, for when red and green durations were equal, the animals more often responded as if red was longer. Perhaps the proper scale for comparison should not be the physical scale based on clock time but instead a psychological scale. Use of a psychological scale that is related to clock time by a power function (with an exponent less than one) produces similar levels of accuracy when the ratios of red to green and green to red are comparable in terms of this scale. The approach of using the psychological scale is reasonable, and previous work has modified scales and durations in light of performance.¹⁶ While an approach that uses a psychological scale as opposed to a physical scale may be reasonable, some caution is necessary. Different experiments have suggested different scales so the present suggestion of a power function conflicts with other suggestions of linear and logarithmic scales.¹⁷ Hasty conclusions should not be made on the basis of one experiment.

Findings of greater control by the second red-green pair are consistent with research involving memory for sequences of events.¹⁸⁻²⁰ When events occur in sequence, performance measures show greatest influence by events at the end of the series. While our results are consistent with animal memory research in some ways, there is one way in which they are not. Research with other tasks as well as duration tasks indicates a lowering of accuracy as the time separating a stimulus and choice increases.²¹ These results imply that accuracy in our task would decrease as the total time of the duration series increased since increases would produce a longer delay between the early durations and a choice. However, findings of similar levels of accuracy over different ranges of durations raise questions about the way in which memory should be considered in this situation.¹⁵

DISCUSSION

At the outset of this paper, a distinction was made between sensory and perceptual frameworks for considering temporal discriminations in animals. Two experiments were reported which grew out of considerations of temporal discrimination from a perceptual approach. The experiments are only preliminary steps; more research is necessary. As one example, pointing in a different direction, Dreyfus and Stubbs trained pigeons to discriminate fixed from variable sequences of events when one choice was reinforced if preceded by a fixed sequence while a second choice was reinforced if preceded by a variable sequence.²² The events were response-produced stimuli rather than time-dependent stimuli, but the results indicate that it should be possible to train pigeons to discriminate different patterns of temporal events.²¹ Collectively the experiments argue for a research program rooted in the perceptual framework. Different stimuli should be used: not just lights, sounds, shocks and food, but stimuli that change in a variety of ways. Different tasks should be used that tap the various processes like those uncovered with human time perception research. And,

experiments should be performed that focus not just on the discrimination of simple duration, but on the discrimination of events with different temporal structure and in different temporal contexts. The study of animal learning has long ignored the findings of perception. Animal time perception provides an opportunity for redressing this neglect by taking advantage of knowledge gained from perception.

REFERENCES

1. RICHELLE, M. & H. LEJEUNE. 1980. *Time in Animal Behavior*. Pergamon. New York, NY.
2. STUBBS, D. A. 1979. Temporal discrimination and psychophysics. *In* *Advances in the Analysis of Behavior: Reinforcement and the Organization of Behavior*, Vol 1. M. D. Zeiler & P. Harzem, Eds.: 341-369. Wiley. Chichester, England.
3. GIBBON, J. 1977. Scalar expectancy theory and Weber's Law in animal timing. *Psychol. Rev.* 84: 279-325.
4. PLATT, J. R. 1979. Temporal differentiation and the psychophysics of time. *In* *Advances in the Analysis of Behavior: Reinforcement and the Organization of Behavior*, Vol 1. M. D. Zeiler & P. Harzem, Eds.: 1-29. John Wiley. Chichester, England.
5. CHURCH, R. M. 1978. The internal clock. *In* *Cognitive Processes in Animal Behavior*. S. H. Hulse, H. Fowler & W. K. Honig, Eds.: 277-310. Lawrence Erlbaum. Hillsdale, NJ.
6. ROBERTS, S. 1981. Isolation of an internal clock. *J. Exp. Psychol. Animal Behav. Proc.* 7: 242-268.
7. GIBSON, J. J. 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin. Boston, MA.
8. GIBSON, J. J. 1975. Events are perceivable but time is not. *In* *The Study of Time: II*. J. T. Fraser & N. Lawrence, Eds.: 295-310. Springer-Verlag. New York, NY.
9. GOLDSTONE, S. & W. T. LHAMON. 1971. Levels of cognitive functioning and the auditory visual differences in human timing behavior. *In* *Adaptation Level Theory: A Symposium*. M. H. Appley, Ed.: 263-280. Academic Press. New York, NY.
10. STUBBS, D. A. 1976. Scaling of stimulus duration by pigeons. *J. Exp. Anal. Behav.* 2: 15-25.
11. GOLDSTONE, S. & W. T. LHAMON. 1974. Studies of the auditory-visual differences in human time judgments: I. Sounds are judged longer than lights. *Percept. Mot. Skills* 39: 63-82.
12. MANTANUS, H. 1981. Empty and filled interval discrimination by pigeons. *Behav. Anal. Let.* 1: 217-224.
13. SPETCH, M. L. & D. M. WILKIE. 1981. Duration discrimination is better with food access as the signal than with light as the signal. *Learn. Motiv.* 12: 40-64.
14. STUBBS, D. A., S. J. VAUTIN, H. M. REID & D. L. DELEHANTY. 1978. Discriminative functions of schedule stimuli and memory: A combination of schedule and choice procedures. *J. Exper. Anal. Behav.* 29: 167-180.
15. FETTERMAN, J. G. & L. R. DREYFUS. Duration discrimination: A pair-comparison procedure. *In* *Quantitative Analyses of Behavior: The Effect of Delay and Intervening Events on Reinforcement Value*. M. L. Commons, J. A. Nevin & H. Rachlin, Eds. Ballinger. Cambridge, MA. In preparation.
16. STUBBS, A. 1968. The discrimination of stimulus duration by pigeons. *J. Exp. Anal. Behav.* 11: 223-258.
17. GIBBON, J. 1981. Two kinds of ambiguity in the study of psychological time. *In* *Quantitative Analyses of Behavior: Discriminative Properties of Reinforcement Schedules*, Vol. 1. M. L. Commons & J. A. Nevin, Eds.: 157-189.
18. SHIMP, C. P. 1976. Short-term memory in the pigeon: Relative recency. *J. Exp. Anal. Behav.* 25: 55-61.
19. WEISMAN, R. G., E. A. WASSERMAN, P. W. DODD & M. B. LAREW. 1980. Representation and retention of two-event sequences in pigeon. *J. Exper. Psychol. Anim. Behav. Proc.* 6: 300-313.

20. HONIG, W. K. 1981. Working memory and the temporal map. *In* Information Processing in Animals: Memory Mechanisms. N. E. Spear & R. R. Miller, Eds. Lawrence Erlbaum. Hillsdale, NJ.
21. CHURCH, R. M. 1980. Short-term memory for time intervals. *Learn. Motiv.* **11**: 208-219.
22. DREYFUS, L. R. & D. A. STUBBS. 1979. Discrimination of a dynamic stimulus. Paper presented at the American Psychological Association, New York.
23. HULSE, S. H., J. HUMPAL & J. CYNX. 1984. Processing of rhythmic sound structures by birds. This volume.

Subjective Duration in Rats: The Psychophysical Function^a

HANNES EISLER

*Department of Psychology
University of Stockholm
S-106 91 Stockholm, Sweden*

Rats are like humans, particularly with respect to time perception. Their psychophysical function is likewise a power function of the form

$$\psi = \alpha(\phi - \phi_0)^{\beta}, \quad (1)$$

where ψ refers to subjective and ϕ to physical duration.

Consider FIGURE 1. The two upper panels describe subjective duration as functions of physical duration. One derives from an experiment with a human observer, the other from rat No. 3. The agreement of the black points with the curves indicates the goodness of fit. Both psychophysical functions build on the parallel-clock model and are derived from duration reproduction data.

Let me briefly recapitulate the parallel-clock model. A thorough description will be found in Eisler,¹ with further clarifications in a subsequent paper.² In the present duration reproduction task, the observer is presented with the standard duration, which is indicated by a sound. After a short interruption, the sound starts again and is terminated by a microswitch (or a lever) being pressed when he, she, or it experiences the second duration as equal to the first. According to the parallel-clock model, the observer deals with two subjective quantities, namely, the subjective durations corresponding to (1) the total duration from the start of the standard to the end of the reproduction ($\phi_T = \phi_S + \phi_V$), and (2) the reproduction (ϕ_V), that is, the variable response duration determined by the observer (ϕ denotes physical time in sec, and the subscripts S , V , and T denote standard, variable, and total, respectively). These subjective durations are assumed to be accumulated in two separate sensory registers ("clocks"), hence the name parallel-clock model. The second duration is experienced as equal to the first when the difference between the total subjective duration ψ_T and the variable subjective duration ψ_V equals the variable subjective duration:

$$\psi_T - \psi_V = \psi_V. \quad (2)$$

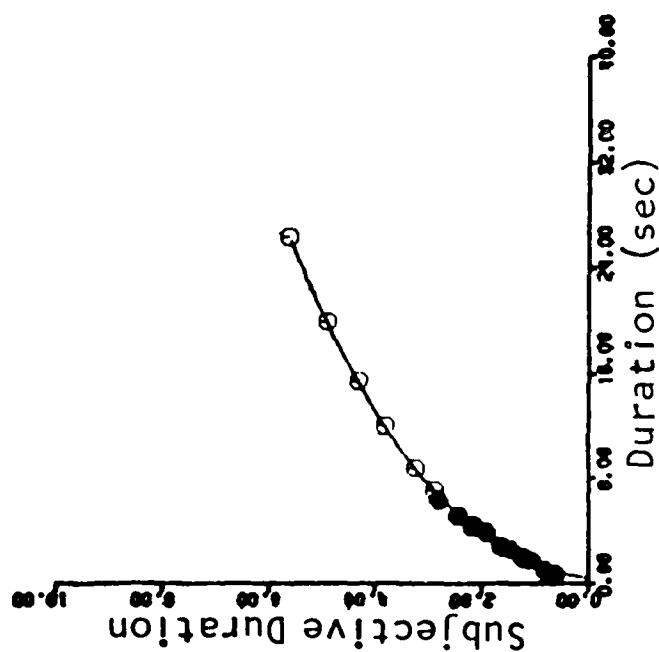
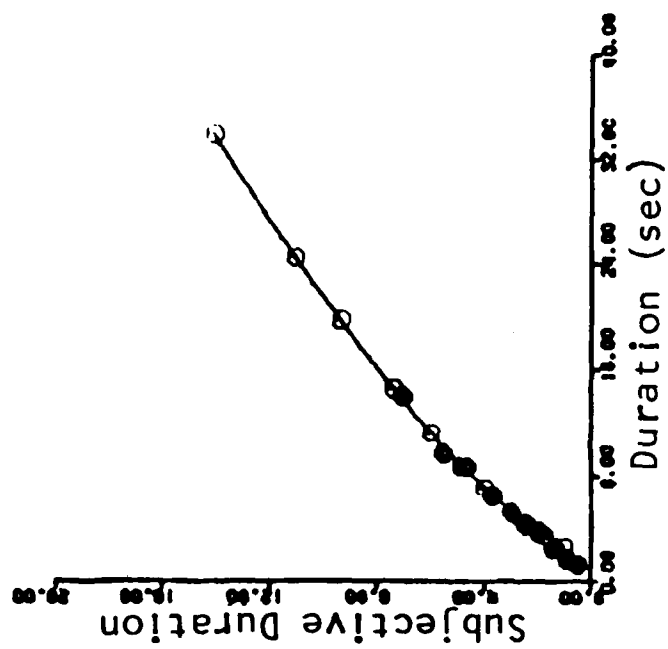
FIGURE 2 should make this clear.

Rearranging Equation 2 shows that $\psi_V = \frac{1}{2} \psi_T$, and inserting Equation 1 yields the following linear relation between ϕ_V and ϕ_T :

$$\phi_V = (\frac{1}{2})^{1/\beta} \phi_T + [1 - (\frac{1}{2})^{1/\beta}] \phi_0. \quad (3)$$

With enough points (standard durations) to fit the straight line, Equation 3 allows computation of the parameters of the power function from the slope and intercept. In

^aThis work was supported by the Swedish Council for Research in the Humanities and Social Sciences. The data treatment was carried out at the Computer Center of the University of Leiden, the Netherlands, during my stay as a fellow at the Netherlands Institute for Advanced Study in the Humanities and Social Sciences at Wassenaar, the Netherlands.



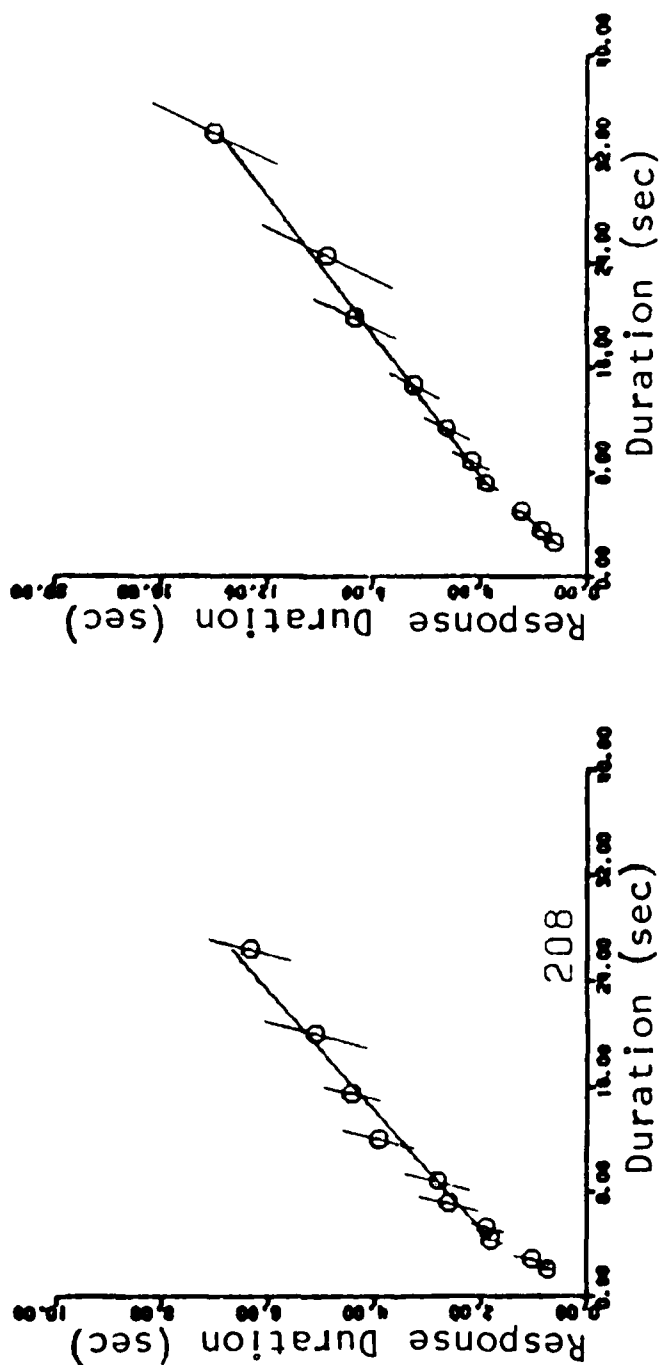


FIGURE 1. Duration reproductions from a rat (*left*) and a human observer (*right*). The *upper panels* describe the psychophysical power function, subjective versus physical duration. *Black dots* indicate goodness of fit. *Lower panels* show plots of response duration ϕ_v versus total duration ϕ_T , together with the fitted straight lines. The *slanted lines* around the points are standard deviations.

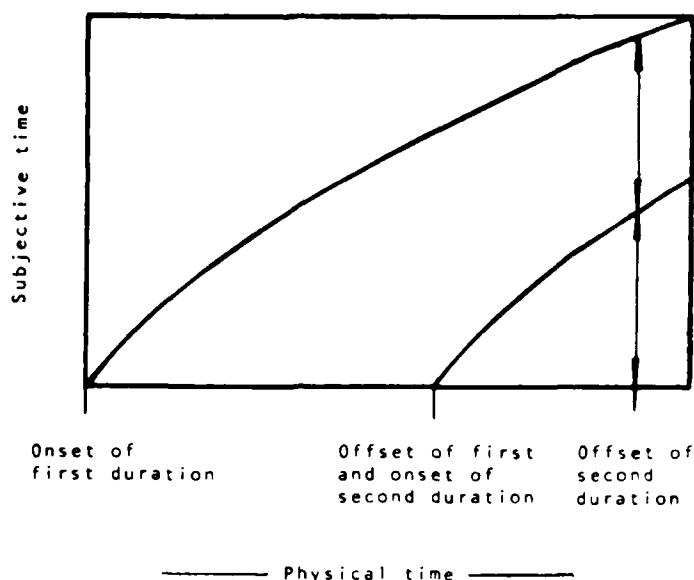


FIGURE 2. Duration reproduction according to the parallel-clock model. Subjective duration is plotted versus physical duration for the total duration (*larger curve*) and the response duration (*smaller curve*). At the point in time when the difference between these two subjective durations (*upper arrow*) equals the subjective response duration (*lower arrow*), the observer reports equality between standard and response duration by pressing the lever (rats) or the microswitch (humans).

this way, application of the parallel-clock model allows one to determine, for example, the exponent β from duration reproduction data. Later I shall describe the data treatment in more detail.

Communication with students of psychology, the data from one of whom I just showed, is comparatively easy, at least compared with communication with rats. But before I go into the details of the experiment, describing the learning procedure that has to replace verbal communication with humans, I want to mention two pitfalls, the avoidance of which probably was a necessary condition for the comparatively successful outcome of my experiment, as opposed to those of Reynolds³ or Mandell and Atak,⁴ for example. The first pitfall is the coupling of two temporal requirements, as when reinforcement after a temporal task is given by a variable-interval schedule instead of using continuous reinforcement (compare the viewpoints of Stubbs⁵).

The other pitfall is the ambition that rats should reproduce time veridically, that is, reinforcing responding within a small band of durations around the standard duration. We know from experiments with humans that reproductions often lie far below the standards: Why should those of rats be closer?

METHODS

Subjects

The subjects were eight male blackhooded rats, obtained 51 days after birth and aged about 1 year and 4 months at the termination of the experiment. They were kept at 80% of their free-feeding weight (compared with a control group of four animals of

the same age) in individual plastic cages in a room without windows in which the day-night light cycle was reversed, so that they were taken to the experimental chamber from a dark room.

Apparatus

The experiment took place in a Skinner box, endowed with a lever and a small loudspeaker, the distorted tone of which indicated the durations. Reinforcement consisted of 5-sec access to sugar-sweetened diluted condensed milk delivered by means of a small movable spoon. I do not want to describe the apparatus in detail here; its function will become clear in the following descriptions of the procedure. Stimuli and contingencies were controlled by means of a punched Mylar tape closed to a loop; the rats' behavior was registered both on paper tape and counters. Here we are concerned only with the response durations ϕ_V .

Learning to Reproduce Durations

The experiment consisted of three phases. The first was just magazine training. In the third phase the subject was presented with a standard duration, followed by a short pause of 300 msec, after which the tone started again until it was terminated by a lever press. The object of the second phase was to teach the rats to attend to this short pause, that is, to an interruption of the tone. Before going into phase 2 in detail, I want to describe procedural features common to phases 2 and 3. Sessions were daily (weekdays and holidays alike) and mostly lasted 1 hour. Each session consisted of cycles, and each cycle of four or five parts (FIG. 3). A cycle begins with a long pause of 30 sec, with no tone and during which a lever press is of no consequence. The next part is the standard duration, indicated by the tone, and thereafter the short pause of 300 msec. Then the tone comes on again and is terminated by the rat's lever press. If the rat's previous

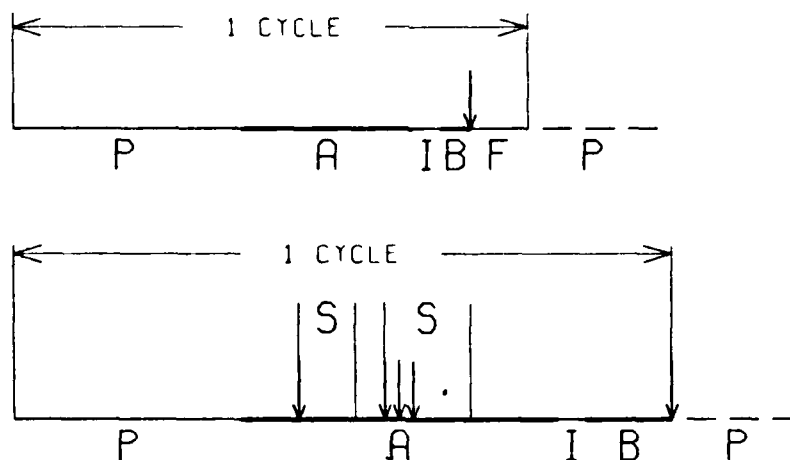


FIGURE 3. Cycle with correct (*upper panel*) and incorrect behavior (*lower panel*) during phase 2. The *thick lines* indicate the ongoing tone, and *vertical arrows* denote lever presses. P = the long pause; A = the presented standard duration, in the *lower panel* lengthened by the punishment time; S; I = the interruption of the tone; B = the latency between the offset of the interruption and the following lever press; F = the feeding period.

behavior during the cycle has been correct (see below), the cycle ends with the feeding period, before the long pause of the next cycle starts. Otherwise the long pause follows directly upon the lever press that terminates the tone. There were ten different standard durations, logarithmically spaced between 1.3 and 20 sec, presented in a pseudo-random series in a block of 30. That is, after 30 cycles, during which every standard duration occurred three times, the whole series was repeated by means of the closed loop of the punched tape. The ten standard durations are found in the second row of TABLE 1.

In phase 2 the rats had to learn to attend to the interruption of the tone. This was achieved by rewarding them for not pressing the lever during the tone before the interruption. A lever press during the standard A (the first tone) lengthened that tone by 4 sec (to avoid superstitious chaining) and reward was withheld, so that the lever press during the second tone was followed by the long pause immediately. If no lever presses occurred during the first tone, the lever press that terminated the second tone was followed by the feeding period before the long pause of the next cycle. Phase 2 was concluded after 68 sessions, at which time the performance of the rats had stabilized at a satisfactory level. There was only a single completely error-free session, and this was achieved by rat No. 3.

In phase 3 the contingencies of lever pressing during the first tone were removed. Instead, the animals were rewarded for pressing the lever during the second tone if the duration of this tone was between bounds that depended on the particular standard duration, irrespective of possible lever presses during the first tone. I made use of two

TABLE 1. The Ten Standard Durations with Lower and Upper Bounds (sec)

Lower bounds	0.35	0.65	1.00	1.40	1.90	2.85	3.85	4.85	6.20	7.50
Standard durations	1.3	1.8	2.5	3.3	4.5	6.0	8.1	11.0	14.8	20.0
Upper bounds	1.30	1.80	2.80	3.80	4.80	6.10	7.40	11.50	15.20	21.00

sets of bounds. The first set consisted of half the standard duration as the lower and 1.5 times the standard duration as the upper bound, and was used for four of the eight rats. The second set, listed in TABLE 1, was used for the remaining four. For these latter bounds, the data obtained from the duration reproduction experiment with humans¹ were used as a guideline. (Originally, I regarded the experiment with humans as a pilot study for the rat experiment.) As the lower bound a value was chosen close to the shortest reproduction given in any of the six trials by any of the 12 observers for that particular standard. The upper bound was chosen similarly. For both bounds, but particularly for the upper, restrictions had to be imposed in order to minimize the overlap between adjacent bands.

Naturally, learning to reproduce durations had to proceed slowly, step by step. Thus, the animals were not exposed to all ten standard durations from the very start of phase 3. The first standard duration used was the second shortest (that is, 1.8 sec), for which a correct reproduction, in accordance with what was said above, had to lie either between 0.9 and 2.7, or between 0.65 and 1.8 sec. These bands have a small overlap with the latencies (B in FIGURE 3) from phase 2, so that the longer of these latencies were reinforced. After a number of sessions with 1.8 sec, the procedure was repeated with 6 and 5 sec, respectively. Then the demands were stepped up further by presenting both 1.8- and 6- (or 5-) sec durations in random order in the same session. Before all ten durations were used, the rats had to work with 20-sec durations, either alone, or in combination with other durations. Which stimuli to choose for a session was to a

certain extent a trial-and-error procedure and depended on the individual rat's performance. Accordingly, the rats had different learning histories.

The experiment was terminated when the performance of the rats deteriorated. I have three possible explanations for this deterioration: (1) The clock controlling the light in the animal housing room stopped working, so that it was dark there at all times. This was discovered only after the termination of the experiment. (2) The rats may have reverted to more "normal" behavior (compare the work of Breland and Breland⁵). (3) Senility may have occurred in the rats.

Data Treatment

The data from each session were dealt with separately. Every standard duration was presented between six and nine times during a session (a result of using a block of 30 for the ten standards). The means of the reproductions for every standard duration were calculated.

However, rats are no better than humans. Rats are occasionally subject to mistakes in their reproductions, which could be ascribed, by analogy with those in humans, to anticipatory errors and slackened attention. That is, there were a few lever presses almost directly after the short pause, and a few very late ones, compared to the rest, for a given standard duration. The rule of thumb I used was to remove reproduction values whenever, after rank ordering, the ratio between the second lowest and the lowest value exceeded 1.7, and in like manner for the highest and second highest. In Session 208 of rat No. 3, for instance, which is the one shown in FIGURE 1, of the 79 reproduced durations, three were excluded from the data treatment. Somewhat surprisingly, the scatter, measured as standard deviation (S.D.), is roughly the same for the human and animal subjects, as indicated by the sloped lines through the data points in the lower panels of FIGURE 1. (Their inclination is due to having fallible values on both axes.)

The rest of the calculations are somewhat more complicated than indicated by Equation 3. Both human and animal data show a break or discontinuity in their psychophysical function, entailing more than one straight line when ϕ_r is plotted against ϕ_s (see FIGURE 1, lower panels). These breaks are described in detail in Eisler,¹ and it is interesting to find them again in the rat data. A break indicates a change in one or two of the parameters in Equation 1, namely, in α , ϕ_0 , or both.

The lines were fitted by the method of weighted least squares, with the reciprocals of the variances of the points as weights. This was recommendable because of lack of homoscedasticity, one cause of which is a tendency for the standard deviations to increase with the ϕ values, that is, to conform to Weber's law, which seems to hold approximately. The other cause is the division of the psychophysical function into two segments on either side of the break. Near the break the psychophysical function sometimes is not defined, so that the subjective duration corresponding to the abscissa at the break can oscillate between the two segments. This shows up as reproduction values that sometimes seem to belong to one, and sometimes to the other straight line when ϕ is close to the transition point. The implication is that the mean is misleading. On the other hand, assigning single values to the one or the other straight line according to appearance is too arbitrary. When using weighted least squares, however, the impact of such a point on the fit is strongly reduced by its large variance.

Unfortunately, often the data are not so clear-cut that one can be completely certain where the break lies. Therefore I carried out the fitting for every data set with the break positioned at all possible values and chose the parameter set that yielded the least sum of squared deviations. About 20 such runs were carried out for every data set.

RESULTS

Of the eight experimental animals, only two "made it," both belonging to the group with bounds based on the experiment with humans, the group with lowest lower bounds. By "made it" I mean that there was at least one session that fulfilled the following two criteria: (1) 50 "correct" reproductions (a session was shut off after 50 feedings) and (2) at least one "correct" reproduction of every one of the ten durations. The reason for the first requirement was two-fold: good performance (about 60-80% correct cycles) and enough reproduction values for each standard duration to obtain a fairly stable mean. The second requirement excluded the "cheating" animals, who tended to reproduce only the shorter standards correctly and press the level much too soon after the longer ones.

All in all, this result can hardly be considered brilliant, with only two out of eight rats furnishing usable data, but in view of the communicative difficulties, it may be regarded as acceptable. Data from 16 sessions of rat No. 3 and from two sessions of rat No. 10 could be used.

By now I have investigated six sessions with rat No. 3 and one with rat No. 10. The exponents β for the first rat varied between 0.43 and 0.70, with a mean of 0.55 and a median of 0.52. The exponent for rat No. 10 was 0.45.

DISCUSSION

As mentioned before, six of the eight experimental rats never learned their task and one of the remaining two produced only two acceptable sessions. The most obvious reason is that communication failed. The data for most of the rats seem to indicate that they, rather than reproducing durations, categorized the standards into a few classes, each eliciting a corresponding response duration. An exponent as low as 0.5 implies that, for a given time interval of middle or long duration, most of the temporal experience is crammed into its first portion,⁶ so that the category "long durations" becomes quite wide. From this it follows that the same response duration is given over a wide range of the higher standards. This interpretation merges with the description of the "cheating" animals given above. It might be worth noting that rat No. 3, the successful one, also was best in phase 2 of the experiment, suggesting individual differences in the sort of ability required for good performance in the present experiment. On the other hand, rat No. 3 was the only one trained during phase 3 with a tape comprising the standards 1.8, 5, and 20 sec, in which the block of 30 consisted of 16 presentations of 20 sec and only seven of the other two standards. So I might just have hit upon a particularly efficient training procedure.

Several conclusions can be drawn from the study reported here. I would like to separate them into three problem areas: (i) procedures for the study of psychophysics in animals, (ii) experienced duration in rats, and (iii) scaling.

Procedures in Animal Psychophysics

In many experiments it has been demonstrated that animals can be brought under the control of temporal stimuli. For an overview see, for example, the work of Stubbs.⁷ However, unlike the present experiment, only one or two durations or classes of durations have been used. Furthermore, the present experiment combines stimulus discrimination and response differentiation in the same task, which has been more the exception than the rule previously. It is quite clear that a rather finely graded discrimination, as well as differentiation, has been achieved. However, the task was

extremely difficult, as can be concluded from the fact that even the best rat's behavior never stabilized in the sense that performance was reliable from session to session. The 16 sessions with data fit for use were interspersed between sessions 183 and 252 of phase 3.

Experienced Duration in Rats

While I obtained a β value of about 0.5, previous attempts to scale subjective time in animals in terms of power functions resulted in exponents that also characterize results in human adults,⁶ that is, a value close to unity.⁶ The reason for this latter value of the exponent probably lies in theoretical confusion regarding the data obtained and the internal representation of duration, as explained in detail by Platt.⁸ Considering that the β value for children is much lower than for adult humans,⁶ the present result gains in plausibility. It is likely that for children, as well as for rats, what happens "now" is more important than what happens later on, that is, that for a given longer time interval, experienced duration is concentrated to its first part (see above).

Scaling

From the scaling aspect it is worth emphasizing that the parallel-clock model seems to work as well with animal as with human duration reproduction data. It is remarkable that the data are so similar, as FIGURE 1 demonstrates, including the break in the psychophysical function. Since this is the third data set accommodating the parallel-clock model (the second are duration discrimination data⁹), the model can be considered well supported. The outcome of the present investigation strengthens the evidence for Stevens' power law (Equation 1) as an internal representation for duration and generalizes it, to my mind convincingly, to animals, or, at least, rats.

I stated at the outset that rats are like humans, at least regarding time perception. Perhaps it would have been more correct to claim that they are like children in that respect.

REFERENCES

1. EISLER, H. 1975. Subjective duration and psychophysics. *Psychol. Rev.* **82**: 429-450.
2. EISLER, H. 1981. The parallel-clock model: Replies to critics and criticisms. *Percept. Psychophys.* **29**: 516-520.
3. REYNOLDS, G. S. 1966. Discrimination and emission of temporal intervals by pigeons. *J. Exp. Anal. Behav.* **9**: 65-68.
4. MANDELL, C. & J. R. ATAK. 1982. Temporal reproduction in the rat. *Behav. Anal. Lett.* **2**: 141-151.
5. BRELAND, K. & M. BRELAND. 1961. The misbehavior of organisms. *Am. Psychol.* **17**: 681-684.
6. EISLER, H. 1976. Experiments on subjective duration 1868-1975: A collection of power function exponents. *Psychol. Bull.* **83**: 1154-1171.
7. STUBBS, D. A. 1979. Temporal discrimination and psychophysics. In *Advances in Analysis of Behaviour. Reinforcement and the Organization of Behaviour*. M. D. Zeiler & P. Harzem, Eds. Vol. 1: 341-369. Wiley, New York, NY.
8. PLATT, J. R. 1979. Temporal differentiation and the psychophysics of time. In *Advances in Analysis of Behaviour. Reinforcement and the Organization of Behaviour*. M. D. Zeiler & P. Harzem, Eds. Vol. 1: 1-29. Wiley, New York, NY.
9. EISLER, H. 1981. Applicability of the parallel-clock model to duration discrimination. *Percept. Psychophys.* **29**: 225-233.

Scalar Timing in Memory

JOHN GIBBON

*New York State Psychiatric Institute
New York, New York 10032; and
Department of Psychology
Columbia University
New York, New York 10027*

RUSSELL M. CHURCH AND WARREN H. MECK

*Department of Psychology
Brown University
Providence, Rhode Island 02912*

INTRODUCTION

A recent report of ours¹ proposed an information-processing account of temporal generalization. The account posited a clock process, which was the basic time measurement device, and working and reference memory for storing the output of the clock either temporarily or relatively permanently. Records of time intervals in working and reference memory were then compared using a binary decision process, which dictated responding or not responding. The analysis concentrated on a relativistic Weber's law property of the data from temporal generalization, and the constraints this property imposed on sources of variance in the information-processing stages. Our purpose here is to summarize that work and generalize the model in two ways: First we consider several sources of variance operating simultaneously. The original analysis demonstrated that if only one source of variance is present, it must be a scalar source, that is, it must result in a variable memory for which variance increases with the square of the mean.² In the generalized account proposed here, we will develop the conclusion that scalar sources dominate in some time ranges, while other sources may dominate in others. These ideas are then applied to two additional timing tasks with different characteristics.

TEMPORAL GENERALIZATION

The reference experiment for temporal generalization is a straightforward task with rats that was reported earlier.³ The procedure is simple: The houselight in an operant chamber is turned off for a duration of time, T . Then a retractable lever is inserted into the chamber. After a 5-sec opportunity to respond, the lever is withdrawn and a 30-sec intertrial interval begins. If the light-off stimulus lasts for the "correct" duration, say, 4 sec, a response is reinforced with a pellet of food. If it is not 4 sec, no reinforcement is forthcoming.

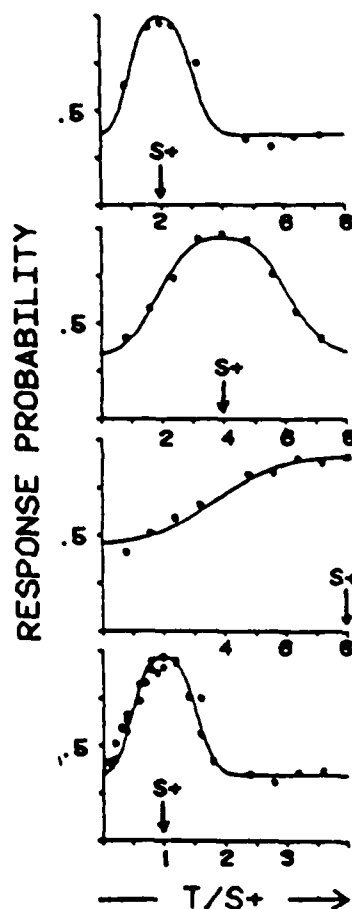
The typical result is that rats come to respond with high probability when the duration is 4 sec, and with decreasing probability for stimuli either longer or shorter than 4 sec.

FIGURE 1 shows how this decline changes with changes in the size of the reinforced duration ($S+$). The probability of responding on the lever for three groups is shown as a function of signal duration when the rewarded duration was either 2 sec (top panel),

4 sec (second panel), or 8 sec (third panel). In each case the maximum response probability is near $S+$, and the spread of responsiveness around $S+$ increases considerably with increases in the $S+$ value.

The way in which the spread is related to the size of $S+$ is shown in the bottom panel, which plots these data on a relative time scale in which signal durations are taken as proportions of the $S+$ duration. The data from the different groups roughly superpose in this metric. This property, which we have called the scalar property, is ubiquitous in animal timing work in the seconds-to-minutes range.² The scalar property we will see exerts strong constraints on admissible sources of variance and on admissible comparison rules for the decision whether to respond or not.

FIGURE 1. Response probability gradients for three groups of rats studied with three different placements of the positive signal. Data points represent median response probability and are replotted in the bottom panel on a relative time scale. (After Church and Gibbon.²)



Our information-processing model for this situation is shown in FIGURE 2. The top row shows the clock process, which includes a pacemaker and a switch for gating pulses to an accumulator in working memory. The pacemaker generates pulses at a mean rate (Δ) that we assume is high relative to the time values (seconds to minutes) that we use in these experiments. The switch, after appropriate training (instructions), gates pulses for a mean duration (D_T) to an accumulator in working memory (second row) when the timing signal is present. The accumulator records and stores the number of pulses (mean of M_T). When, at the end of a given trial, a response is made and reinforced ($T = S+$), the time value recorded in working memory on that trial is stored in a more

permanent reference memory for reinforced values (mean of M_S^*). The third row shows the decision process. A response occurs when a comparator yields a judgment that the current record in working memory for this trial is "close enough" to the reference memory for the reinforced duration to warrant a response.

In FIGURE 3 we show the clock process in more detail. Imagine a pacemaker generating pulses with interpulse intervals, τ , with mean rate $\Lambda = \lambda$. In this figure, the pulses are evenly spaced to indicate no variance. In later analyses we consider variance in these parameters.^a

The pulses are switched into the accumulator by the switch indicated in the middle box. The switch (SW) is assumed to have some latency to close (t_1) after the signal goes on, and some latency to open (t_2) after the signal goes off. Thus, the mean time during which pulses are gated into the accumulator is $D_T = T - T_0$, where T_0 is the expected

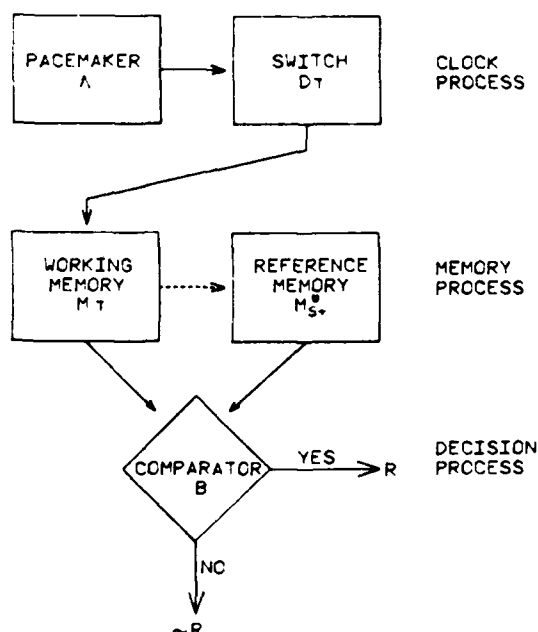


FIGURE 2. Information-processing model. The top row describes a clock process with a pacemaker generating pulses that are switched into an accumulator for working memory (middle row). After reinforcement, working memory contents are stored in a reference memory for later comparison on subsequent trials with the current working memory value.

difference between the latencies to close and open the switch. In principle, T_0 may be negative as well as positive. The graph above the switch shows this linear relation between D_T and T . The intercept is the minimum signal duration below which counts do not register in the accumulator. For signal durations less than this minimum, the switch is opened before it has closed so no pulses are switched into the accumulator. Conversely, if the latency to reopen the switch exceeds the initial latency to close it ($t_2 > t_1$), even a very short signal duration suffices to allow counts to register during the t_2 latency. If t_1 and t_2 were precisely the same, the switch duration would mimic exactly the signal duration. In general, this situation is unlikely, however, and these two latencies constitute one of the sources of variance that we will consider later.

The accumulator in this scheme simply records the number of impulses gated to it. The mean accumulated number, M_T , therefore, is just the impulse rate times the

^aWe generally refer to random variables with lowercase letters and their expectations with the corresponding uppercase letter (for example, $D_T = E(d_T)$). An exception is λ , which can play a dual role in the Poisson pacemaker.

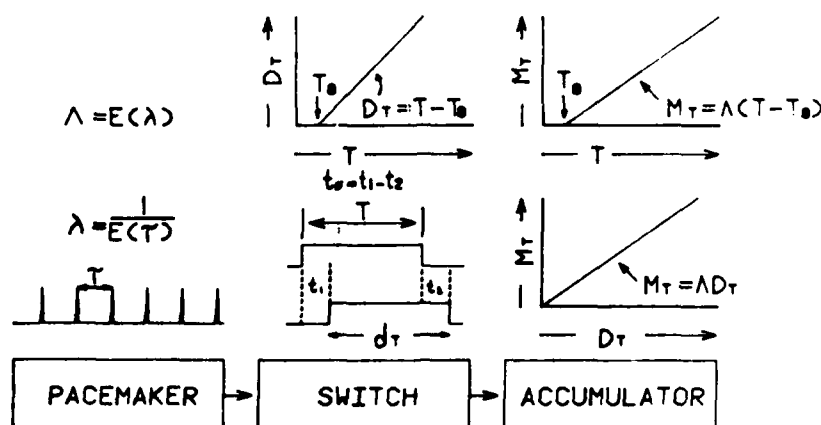


FIGURE 3. Clock process. The pacemaker generates impulses at intervals of τ and these impulses are gated into an accumulator by the switch. The switch has a latency to close at the beginning of the stimulus and to open at the end of the stimulus. The accumulator records the number of impulses gated to it during the interval that the switch is closed. (After Gibbon and Church.¹)

duration that the gate is closed, as shown in the lower graph directly above the accumulator. The upper graph shows the accumulation value as a function of signal duration. The switch introduces the non-zero intercept.

The memory system that we propose here is much simpler than is probably realistic for memory models (for example, that of Heinemann^{4,5}). However, it is quite complicated enough to analyze, even in this oversimplified form. We propose that the working memory directly reflects the accumulated count as shown in the proportional plot above working memory in FIGURE 4 (and in FIGURE 3). When, for a given

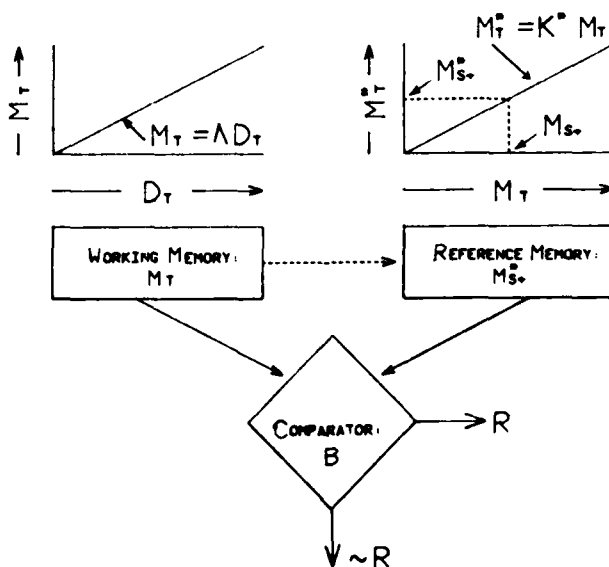


FIGURE 4. Memory process. The accumulator in working memory directly reflects counts gated to it from the pacemaker. On reinforced trials the working memory record for that trial is stored in a more permanent reference memory. (After Gibbon and Church.¹)

condition (series of trials), $S+$ is fixed at some value, and when on a particular trial $S+$ is presented ($T = S+$) and a response is made and reinforced, the value stored in reference memory, m_s^* , is a proportional representation of the value on that trial recorded in working memory (m_T). The comparison, then, on subsequent trials, is between the current value in working memory, m_T , and a stored value in reference memory, m_s^* . The judgment whether or not to respond is based on some kind of comparison between these two values.

Several major features of this account are present in an early proposal of Treisman,⁶ in which a pacemaker, counter, store (memory), and comparator occur. To our knowledge, that was the first model to use a form of scalar timing (equation 11) explicitly in a timing system. Creelman⁷ some years ago also proposed a timing model in which a pacemaker and a counter were involved, but with a Poisson rather than scalar form of variance. These distinctions will come in for more discussion later.

In our earlier analysis¹ we discuss briefly the need for separate lines of evidence to establish the existence or utility of each of the processes in this account.⁸ In the present article we will concentrate on features of the account that may be applied in a similar manner to similar and dissimilar timing tasks.

First, however, a brief summary of our earlier analysis is necessary. We first examined a case in which there was no variance in the system whatever, so that the values in the accumulator in working memory and in reference memory were all one-to-one with the appropriate signal durations.

Two response rules were examined, one of which compared the current time value in working memory with the remembered time value in reference memory, and dictated a response when these two values were "close enough," that is, when the absolute discrepancy between them lay below a certain threshold. The absolute rule for the no-variance case is shown in FIGURE 5 in the top row. The ticks on the abscissa in the upper left panel indicate respectively, T_0 , S , and $2S$, corresponding to the minimum signal duration, a given $S+$ duration, and a second $S+$ duration set at double the first. These might be appropriate, say, to the 2-sec and 4-sec $S+$ conditions in FIGURE 1. The absolute difference between M_s^* and M_T is shown decreasing to 0 at S and increasing beyond it as the solid line function. The positive diagonal hatching shows the acceptance region within which responses are required when the absolute discrepancy falls below the threshold value indicated by the horizontal line, B . When the $S+$ value is doubled (at $2S$), the dashed line shows the discrepancy function for this case, and the negative diagonal hatching indicates the acceptance region.

In the middle panel of the top row the response consequences are plotted in real time. In the first case for $S+ = S$, responding occurs whenever T is within an absolute window of S , and the same rule for $S+ = 2S$ produces the same spread around the reinforced $S+$ value. In the upper right panel these step functions for response probability are replotted in relative time, as a function of $T/S+$. The efficiency of the absolute rule is seen to increase considerably in relative time as $S+$ is increased. The spread in relative time around $S+ = 2S$ is much smaller than that around $S+ = S$ in this plot. This plot is comparable to that in the bottom panel of FIGURE 1. It shows that an absolute comparison rule is untenable under these assumptions.

In the bottom row of FIGURE 5, we show an alternative response rule, which compares the absolute discrepancy between the current time and the remembered reinforced time, to the remembered reinforced time. This relative rule requires a response whenever

$$\left| \frac{M_{S+}^* - M_T}{M_{S+}^*} \right| < B. \quad (1)$$

The subjective discrepancy is taken as a proportion of the reference memory value for reinforcement. In the panel in the lower left, the solid line function for $S+ = S$ is shown decreasing from 1.0 down to 0 and back up for $T > S$ as in the panel above. Now, however, the threshold, B , is a proportion. When the discrepancy is less than this proportion, responding is dictated. Again the acceptance region is indicated by positive diagonal hatching.

The dashed function shows the relative subjective discrepancy for $S+ = 2S$. When $S+$ is doubled, the window size nearly doubles also, as the negative diagonal hatching shows. In the middle panel the response consequences plotted in absolute time show the increasing spread around $S+ = 2S$, and in the right panel the two acceptance regions

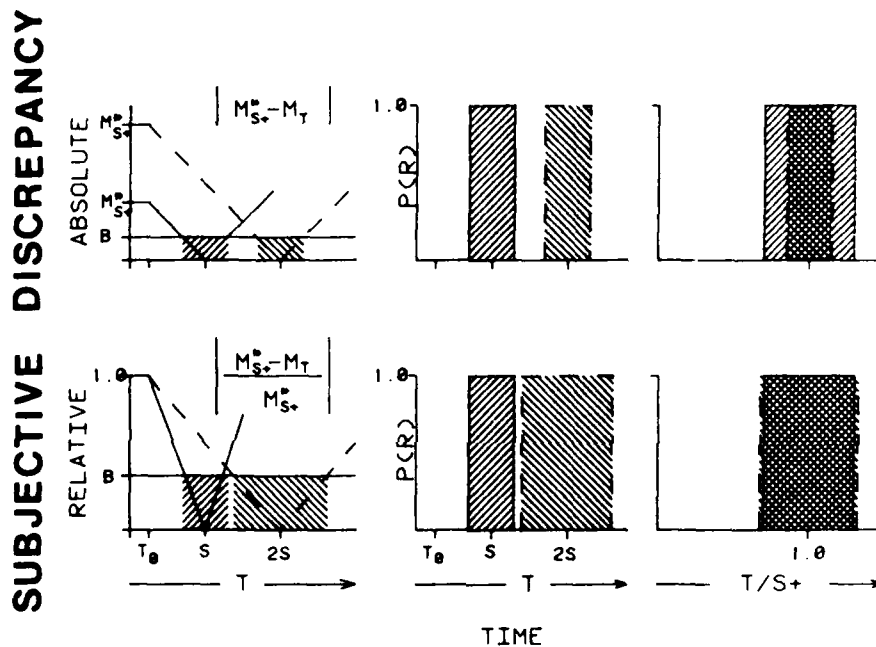


FIGURE 5. Two comparison rules. The *left column* shows subjective discrepancies between working memory and reference memory as a function of signal duration with two different placements of $S+$. The *right two columns* show response probability. The *top row* describes an absolute discrepancy rule, in which the absolute difference between working and reference memory values forms the basis of the response decision. The *bottom row* describes a relative discrepancy rule in which the absolute discrepancy is taken as a proportion of the representation of $S+$. (After Gibbon and Church.¹)

are nearly equivalent when plotted in relative time. They are not precisely equivalent because the T_0 intercept in the accumulation of subjective time counts more heavily when $S+$ is smaller than when it is larger. However, even for T_0 relatively large, the difference is not great. In this example T_0 has been chosen equal to $S/4$.

Thus, even when there is no variance in the timekeeping and mnemonic system, a relative response rule is needed to accommodate the superposition of generalization gradients in relative time.

Of course, the real response gradients are not square waves, and so a no-variance account will not suffice as a realistic description. In our earlier analysis we studied several sources of variance which would generate smooth shoulders on the gradients.

but we found that in no case was the absolute discrepancy rule tenable. Even when variation increased substantially with $S+$, the absolute discrepancy rule is too efficient to accommodate a realistic description of the data as $S+$ is increased.

With the relative discrepancy rule, however, a number of different sources of variance were compatible with both the superposition of the data in relative time and with smooth shoulders around $S+$ (FIG. 1). The analysis showed that at least four distinct sources of variation in perceiving, remembering, and discriminating time intervals were possible. Each of these sources was scalar, that is, multiplicative with the memory for time. Each was associated with different components of the information-processing scheme. Two additional sources of variation in the system are also potential contributors to variability, but these sources were shown to be not feasible as a sole source of variation, were there but one.

One of our current goals is to deepen this analysis by allowing these latter two sources of variation to be imbedded in variation from multiple points in the system. The earlier analysis is summarized with respect to the relative discrepancy rule only. The way in which the two nonscalar sources of variation violate the superposition requirement will be described briefly first.

Switch: Constant

The simplest of these two sources of variation is the variability introduced by varying latency to open and close the switch gating pulses into short-term memory. It is likely that even in highly practiced subjects there is some variation in the perceptual

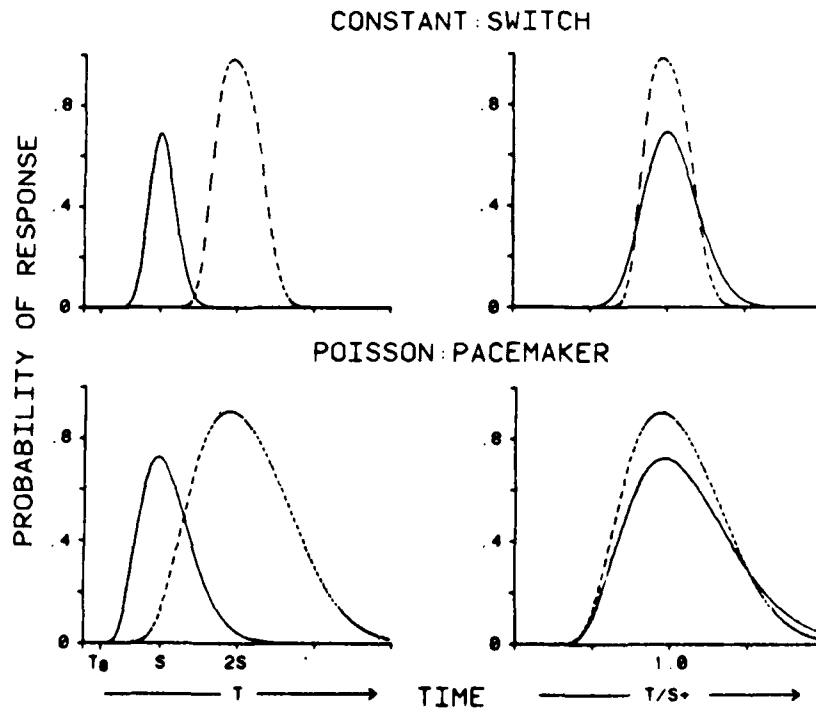
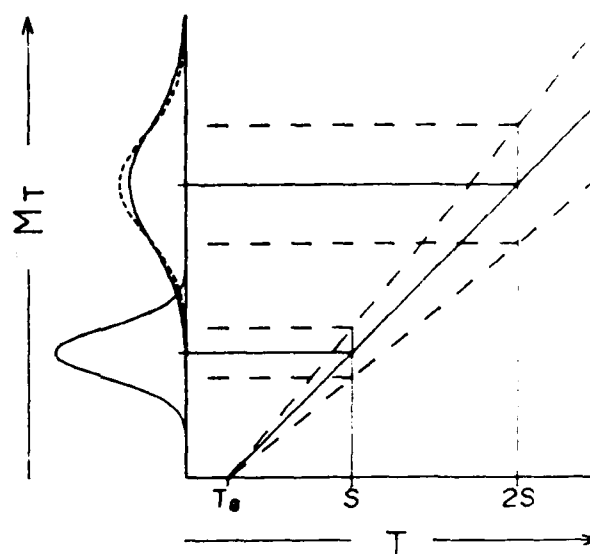


FIGURE 6. Response probability gradients resulting from variation in the switch only (top) or from Poisson variation in the pacemaker only (bottom). At the left gradients are plotted in real time and at the right in relative time.

FIGURE 7. Working memory accumulator. Distributions of counts associated with two different $S+$ durations are shown when the only source of variance is the scalar fluctuation in pacemaker rate. (After Gibbon and Church.¹)



response to external stimulation. A gating mechanism introducing variance of this sort is similar to an early proposal of Kristofferson and his colleagues which argued that in some ranges the timing system might reflect only constant, low variance.^{9,10,11} A realistic form for this latency difference distribution might be most likely back-to-back exponentials (Laplace distribution). A similar alternative is the triangular distribution analyzed by Kristofferson as the result of a difference between two rectangular variates. If the switch is complex, however, and requires the execution of several components in order to open and close, it might approach a normal form. The niceties of distinctions between forms will not concern us in the present account, since we will rely on rather grosser distinctions which are relatively insensitive to distribution form. We assume a normal form in what follows.

The top row of FIGURE 6 shows the effect on response probability of introducing variance in switch latency only. In the upper left panel response gradients are plotted against absolute time for $S+ = S$ and $S+ = 2S$. Increasing $S+$ results both in greater accuracy at $S+$, and in a broader spread around $S+$. In the panel in the upper right, these effects counterbalance each other somewhat in relative time. There the largest discrepancy is at $S+$, while the increase in the spread with increasing $S+$ is seen to be not sufficient to produce superposition in the wings. Constant variance, which does not change with increasing size of the interval being judged, results in too efficient a system at large values of $S+$. Accuracy is increased at the target duration, and the spread around it does not increase sufficiently to accommodate superposition.

Pacemaker: Poisson

A second source of variation in this system is one that has received classical attention in modeling perceptual systems, namely Poisson variation in the discharge of the pacemaker. From a variety of considerations, variance in the interpulse interval in a neural pacemaker ought to follow the Poisson law.^{6,7,12,14} Interpulse intervals are exponentially distributed in a steadily varying stream with intensity $\lambda = \Lambda$. The accumulator in memory is then a Poisson counter, with a variance that increases directly with the mean.

In the lower row of FIGURE 6 response gradients for the Poisson pacemaker source are shown. The gradients in absolute time in the lower left show an increase in accuracy at $S+$ when $S+$ is doubled. However, this increase is not as large as that for the constant variance case, and the increase in spread at $S+ = 2S$ is somewhat broader. Nevertheless, when plotted in relative time on the right, the gradients still deviate substantially from superposition. Thus, this system, which does increase variance with increasing $S+$, does not do so fast enough to accommodate superposition, particularly near $S+$.

These two kinds of variability introduced by different stages of information processing are to be contrasted with two other kinds of variability, both of which are scalar and induce the approximate superposition seen in the data. These are discussed next.

Pacemaker: Scalar

An alternative source of pacemaker variance is a drifting rate. For simplicity we imagine that the time between pulses, τ , is fixed on any trial, but that from trial to trial the pulse rate, $\lambda = 1/\tau$, varies normally around a mean, Λ .^b

In FIGURE 7 the result for the accumulated count in memory is shown for $S+ = S$ and $S+ = 2S$. The rising solid line reflects the mean value of the rate, Λ , comparable to the rising line in the upper right panel in FIGURE 3. The dashed rays emanating from T_0 indicate plus and minus 1 standard deviation of λ . When $S+$ is doubled, the distribution of accumulated counts is nearly a scale transform of the distribution associated with S , hence the name scalar timing. The solid function distribution for $2S$ is generated by our information-processing system with a fixed T_0 . The dashed function which is nearly identical to it is the simple scale transform of the S distribution that would result from strict proportionality, an axis multiplier of 2. The small discrepancy reflects the role of T_0 . Again, T_0 was chosen rather large, equal to $S/4$. Even with an intercept of this size, the multiplicative property dominates the result.

This kind of multiplicative variance in combination with the relative discrepancy rule results in the response gradients shown in the top row of FIGURE 8. In absolute time (on the left), it is clear that accuracy when $S+ = S$ and when it equals $2S$ remains the same, but variance increases considerably. In the proportional plot in the upper right panel, the functions now show near superposition, like the data. Thus, if the only source of variance were a drifting rate in the pacemaker, the smooth shoulders, constant discriminability at the positive value, and rough superposition are accommodated by this scalar source.

Memory: Scalar

In our earlier analysis we also studied the possibility of additional noise introduced when the accumulated counts from the pacemaker are stored in working memory, and again when these records are transferred to relatively permanent storage in reference memory. If both of these storage mechanisms involve a proportional transform, then they too may introduce variability which has the scalar form shown for pacemaker rate

^bA more realistic version might allow both kinds of variation simultaneously. Our later analysis (APPENDIX) allows Poisson variance within trials at a given intensity, λ , which in turn is a random variable that drifts over trials.

variance in the top row of FIGURE 8. The first of these potential sources, that involved in translation to a short-term storage in working memory, will be ignored in what follows here, since with the ratio discrimination rules that we employ, the mean value of the translation process is cancelled, and its variance may be absorbed in the pacemaker variance term. Storage in reference memory, however, plays an important role, and may contribute variance as well as distortion between remembered and current working memory values. Qualitatively, however, these two sources of variability show similar results for the response gradients as long as the reference memory translation is unbiased. Hence we do not treat them separately here.

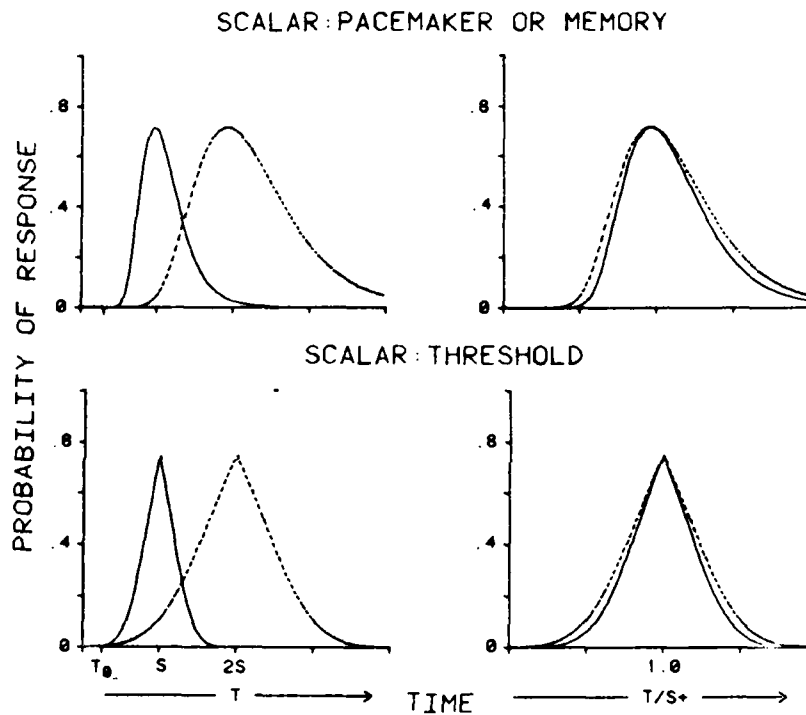


FIGURE 8. Response probability gradients for two scalar sources. The *top row* shows the results of variation in either the pacemaker or the memory constant. The *bottom row* shows the results of variation in the threshold.

Threshold: Scalar

Still another source of variance is realistic for our processing system. This is variance in the relative proximity of T to $S+$ that subjects deem "close enough." One can readily imagine momentary fluctuations in this level. Multiplying through by the norming value in Equation 1, variation in b induces the scalar property on remembered time. The manner in which this property is expressed, however, differs from that for clock or memory variance.

In the bottom row of FIGURE 8 response gradients are plotted assuming a normally distributed threshold about some positive value b . The gradients in the lower left panel are strictly symmetric, have identical accuracy for $S+ = S$ and $S+ = 2S$, and show a prominent discontinuity at the positive value. In the lower right panel the relative time gradients approximately superpose. Thus, this source of variance, like the scalar source

for clock or memory in the upper row, induces superposition in relative time, as the data require. The manner in which it is accomplished, however, differs in two ways: First the upper gradients have a slight, positive skew. The lower gradients in contrast are strictly symmetric. Second, the upper gradients are smooth, roughly bell-shaped around $S+$, while the lower gradients are discontinuous at $T = S+$. The discontinuity results from a feature of our assumptions that is open to question here, namely, a normal form for the threshold distribution that includes negative values. In practical terms this assumes that on some trials the threshold is so conservative that not even a comparison of identical values (no subjective discrepancy) would warrant a response. The defect may be remedied by truncating threshold distributions to be always positive, but the result is equally difficult for theory if this is the only source of variation. This modification induces perfect accuracy at $T = S+$ for all $S+$ values, and the data show clearly otherwise.

These features of the threshold variance account might prove troublesome for a description involving only this source of variance. However, in combination with the clock and memory variability we will see that the symmetry in threshold variability is an important feature of our account.

Simultaneous Sources of Variation: General Case

From the analysis thus far, we were able to conclude that were one forced to pick a single source of variation, it would have to be a scalar source, and probably would have to be located in pacemaker or memory. However, while the smooth character of these gradients looks like that of our temporal generalization gradients, and while our temporal generalization gradients have, as do these, some slight skew with a higher right than left wing, the skew in theory is generally larger than the skew in reality. The data fall somewhere between strict symmetry around $S+$, and the scalar forms for clock or memory variance alone.

This discrepancy leads us to analyze the mixture generated by allowing all of these potential sources of variability to operate at once. In the APPENDIX we pursue this analysis, obtaining an approximation for the underlying random variables in working and reference memory, and their discrimination via a variable threshold in a comparator. The result may be summarized as follows: Allowing random variation in constant, Poisson, and scalar sources results, not surprisingly, in constant, Poisson, and scalar components of variance in the composite variables underlying the spread of the response probability gradients. For reasonable choices of parameter values for the constant and Poisson variance sources, however, the scalar sources dominate variance in the seconds to minutes ranges studied here. This may be seen in FIGURE 9. In the APPENDIX the argument is developed showing that under our assumptions the response probability gradients reflect two random variables, one corresponding to each edge of the acceptance region for our comparison statistics. The standard deviations of the window edge variables grow (nearly) linearly with T and $S+$. In FIGURE 9 the standard deviation of the upper window edge is shown as a function of increasing $T - S+$ signal durations for a variety of choices of pacemaker rate, $\Lambda = 5, 10, 15$, and 20 per sec. The figure assumes a switch with a mean and variance of $T_0 = \sigma_0 = 0.5$ sec, variance in the threshold of $\sigma_b = 0.1$, and a scalar pacemaker or memory coefficient of variation of $\gamma = 0.2$. The point of the figure is that the scalar sources of variance rapidly come to dominate the standard deviation of the window edge. These functions are approximately linear in $S+$ for $S+$ values above about 5 sec (dashed line). Since the ranges we study here are usually well above this level, we cannot say whether significant contributions from a Poisson source and a constant source might not be

present, but masked. The domination by the scalar sources in these mixtures may be used in the future to specify more precisely the range of parameter values for pacemaker rate and switch variance that are possible and yet still accommodate our data. It is clear from the work presented in this volume that the temporal ranges that interface between milliseconds and hours will be important in the future.

PEAK

A procedure that is very similar to temporal generalization in conception and results, but quite different in the manner of assessing response strength, is the "peak" procedure developed by Seth Roberts¹⁵ and studied by him and Church and Meck.¹⁶ It is a modified discrete-trial free-operant fixed-interval schedule. At the onset of a signal, usually a light or a noise, rats may respond on a lever, which occasionally will "pay off" for the next response after $S+$ sec have elapsed from the beginning of the

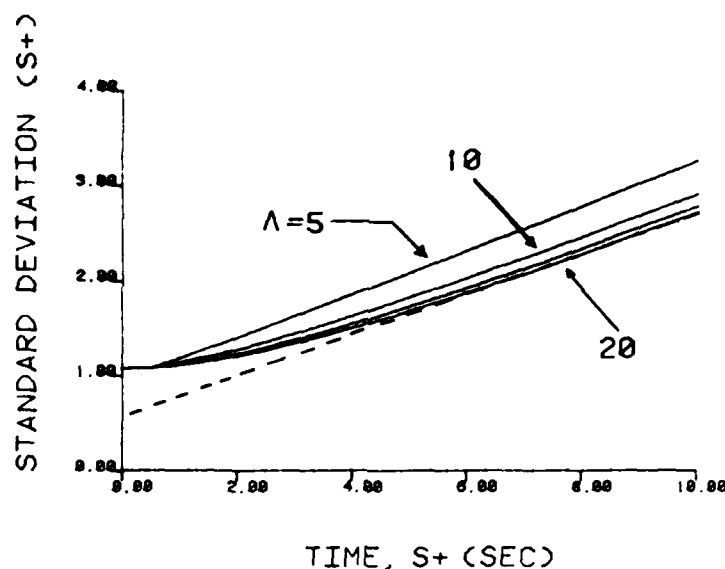


FIGURE 9. Standard deviation of a window-edge decision statistic as a function of the size of $S+$. The parameter is pacemaker rate.

signal. On some trials, however, responding is not reinforced and the signal remains on for a long time. Subjects come to anticipate completion of the $S+$ interval. On trials in which reinforcement is not programmed, they show maximum responding close to that time. Typical results are shown in FIGURE 10. The top row shows responding of two pigeons studied under the peak procedure using a color change on a response key as the $S+$ duration signal. On half of the trials the first response after 15 sec was reinforced. Response rate increases up to a maximum near 15 sec and decreases thereafter.

In the second row, data from a rat studied by Roberts is shown for two different conditions. In the first condition on 50% of the trials the rat was reinforced for responding after 20 sec had elapsed since the onset of an auditory signal. The leftmost function shows the performance. Response rate peaks a little beyond 20 sec and declines at longer times. The broader function on the right was obtained from the same rat under a later condition with $S+ = 40$ sec. These data are quite similar to the

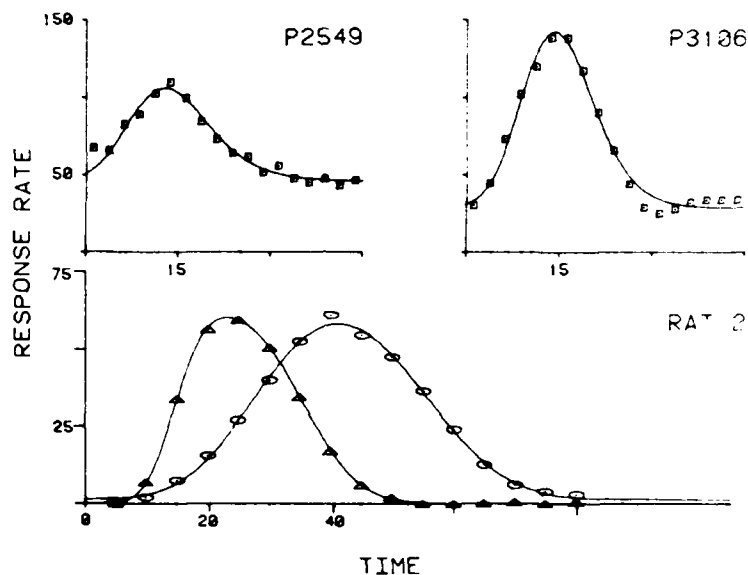


FIGURE 10. Response rate gradients from two pigeons (*top row*) and one rat (*bottom row*) studied under the peak procedure.

temporal generalization data in two respects: Accuracy at the peak time is about the same in both functions, and the spread around the larger $S+$ is approximately proportionally increased. Indeed, the peak procedure might be thought of as a generalization procedure which delivers all possible durations on every unreinforced trial.

Memory Distortion

All of the above cases show some slight displacement of the peak time from the nominal $S+$ value. This is not unusual with the peak procedure. The peak for bird No. 3106 is about 13 or 14 sec and the peak for rat No. 2 studied under $S+ = 20$ sec occurs near 24 sec.

In FIGURE 11 in the left hand column the effect of deviation from $S+$ is shown for the data of rat 2. The top functions show the peak rate as a function of absolute time as in FIGURE 10. In the middle panel, these functions are shown plotted in relative time, $T/S+$, as in FIGURE 1 for temporal generalization. Now superposition is not achieved. The $S+ = 20$ sec curve shows a broader right wing in relative time.

In the model, the information-processing account allows for distortion of the mean in the translation constant between working and reference memory.¹⁷ Working memory is assumed to reflect pacemaker accumulation directly, while the storage into reference memory may introduce both additional variance of its own, and systematic distortion in the mean (upwards or downwards). The translation constant, K^* , required for these data is somewhat above 1.0, hence the failure of superposition.

In the lower left panel, a fit to the peak data has been accomplished and the data are now plotted relative to K^*S+ , that is, relative to the mean of the memory distribution for the reinforced time. When adjusted in this manner, superposition succeeds.

A more extreme example of memory distortion which nevertheless preserves

superposition in the appropriate metric is obtained from an experiment studying two peak procedures in rats simultaneously. A peak procedure value of 10 sec in the presence of a light and 30 sec in the presence of a noise were studied on randomly intermixed trials. Median performance of a group of ten rats is shown in the right hand column of FIGURE 11. In the upper right panel the two peak functions plotted in real time are seen to be centered close to 10 sec and 30 sec, as would be expected from the preceding data. The functions differ somewhat from those of the individual subjects in FIGURE 10. Both functions are fairly flat near the peak, leaving some ambiguity as to just which time value corresponds to the "true maximum." However, even with this flatness it is clear that the 10-sec data lie mostly above 10 sec, while the 30-sec data lie

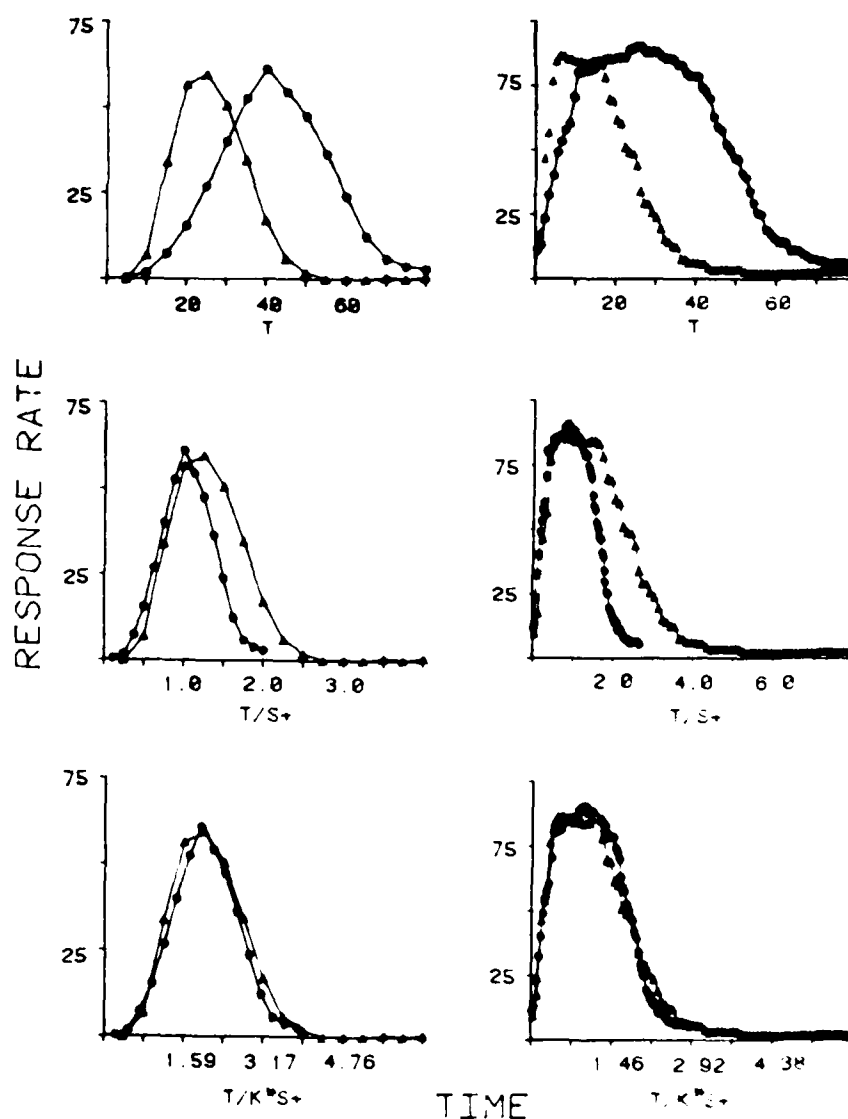


FIGURE 11. Response rate gradients for rat 2 in the left column, and for a group of rats studied under two different $S+$ values on the right. Gradients in the top row are plotted in real time, in the middle row in time relative to $S+$ time, and in the bottom row in time relative to the empirically determined memory representation of $S+$.

mostly below 30 sec. It is as though the subjective representations of the 10-sec and 30-sec $S+$ experiences have become mixed or perhaps mutually attracted in memory.

In the middle panel of FIGURE 11 these data are plotted in relative time, and the systematic distortion is again evident. The 10-sec curve appears broader than the 30-sec curve in relative time, because its "true" $S+$ time is larger than 10 sec and the "true" 30-sec $S+$ time is less than 30 sec. This is similar to the deviation on the left, but more extreme. In the bottom panel, the results of fitting these data, allowing for memory distortion, show clear superposition when plotted in time relative to the memory time ($T/K \cdot S+$).

Thus, the peak procedure provides temporal gradients with a maximum near the time of reinforcement, but deviations appear both idiosyncratically and because of potential interactions in reference memory when more than one time value must be retained. The results of this analysis imply that this sort of distortion is a multiplicative one. The scalar property applies to the memory for time, not necessarily to real time.

TIME LEFT

Our final application of these ideas is to a procedure that differs considerably from the temporal generalization or peak procedures. In a previous report, Gibbon and Church¹⁸ studied choice procedures with rats and pigeons; these procedures were designed to reveal preference for the subjectively shorter of two delays to food, when one of these delays was elapsing. The aim of that work was to examine these choices parametrically at a variety of delays, since parametric data should reveal curvature or linearity in subjective time. The linearity result will not be reviewed here for that implication, but rather some new data from the same procedure under several conditions will be examined for the superposition property.

The procedure studied with pigeons is shown in FIGURE 12. At the beginning of the trial (initial link), two keys are lit with different colors, say white and red, and birds distribute responding across the two keys as time elapses during the trial. At some point (T), the next response produces mutually exclusive (terminal link) consequences on either key. If the next response is the white key, the red key is extinguished, and responding on the white key may continue for the remainder of the white-key interval,

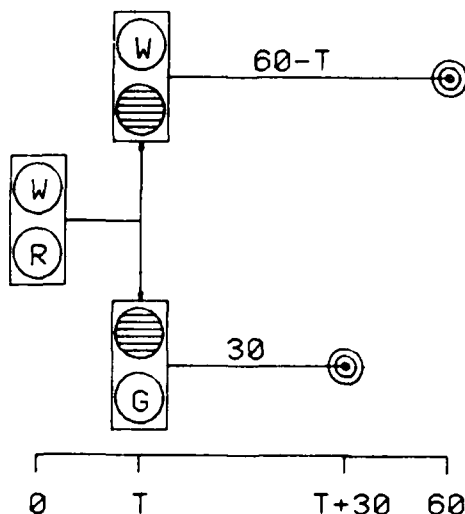


FIGURE 12. "Time left" procedure. Two keys are lit with different colors (white or red) at the beginning of the trial. At some randomly chosen point, T , the next response results in one or the other of two mutually exclusive delays to reinforcement (bull's-eyes).

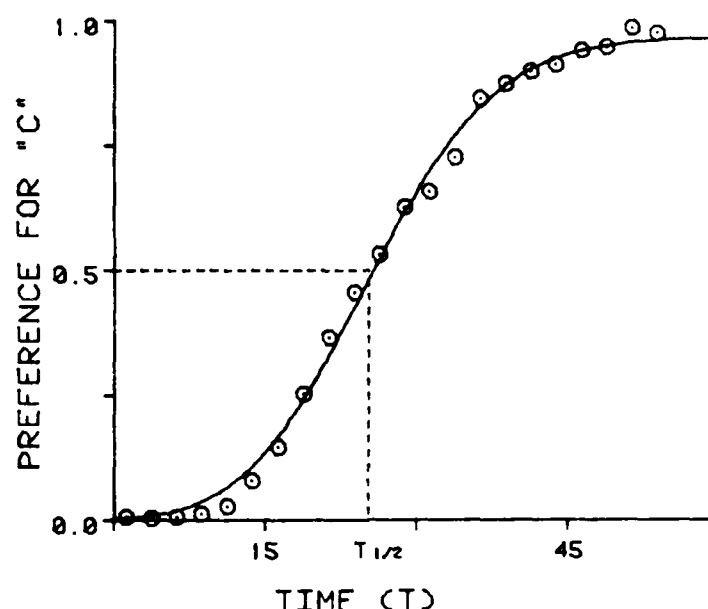


FIGURE 13. Psychometric preference function. Preference, the proportion of "C" choices at successive points during the choice period, increases from favoring "S" at the start of the trial to favoring "C" as the trial elapses. The point of indifference is indicated by the dashed line over $T_{1/2}$.

in this example 60 sec, when reinforcement is made available. Thus, on the white key, reinforcement is available after a total of 60 sec from the beginning of the trial.

If, at the choice point, the next response is to the red key, the white key is extinguished, the red key changes color to green, and responding will be reinforced for pecking the green key after 30 sec have elapsed since the color change. Thus, this delay, called the standard (S), is fixed at 30 sec. The entry times (T) vary from trial to trial.

Early in the trial, it behooves subjects to respond to the red key since the delay to food there is shorter than on the comparison 60-sec interval. However, as the choice period elapses, the time left on the comparison ($C = 60$ sec) side becomes shorter than the 30-sec standard, and now it behooves subjects to respond on the white key. The typical result is shown in FIGURE 13, in which choice responding begins heavily favoring the standard, and at some point in the trial switches over to favor the comparison time-left side as that delay becomes more favorable. The point at which the switch is made, $T_{1/2}$, is a datum of primary interest. If subjective time is linear and subjects are unbiased in their appreciation of this time, then they should switch over to preferring the time-left side of the choice at precisely $C - S$ sec into a C -sec comparison interval—at 30 sec into the 60 sec interval in the above example. In fact, it is common for subjects to switch somewhat before the midpoint value for $C = 2S$, and the bias in favor of the elapsing interval may be due to a preference for a key color paired with primary reinforcement. In the analysis that follows, the crossover point, $T_{1/2}$, plays a fundamental role.

The information-processing account for this procedure has three distinct conditions: the choice period, and each of the mutually exclusive terminal link consequences. The scheme we use to model choice responding is shown in FIGURE 14. When the choice period begins, the switch gates pacemaker pulses into a working memory accumulator for trial time, T . The difference between the memory for current time and the memory

for the comparison (C) interval represents the subjective time left on the comparison side. This remaining time is compared in the comparator with memory for the time "left" on the standard side—the standard delay to food were the terminal link to come on immediately.

As time accumulates during the trial in working memory, the remaining time on the comparison side is continually updated, and compared with memory for the standard delay. At the outset of the trial, the standard delay appears shorter than the remaining time on the comparison delay and hence preference for the comparison interval is low. At some point subjects cross over and prefer the remaining time on the comparison side, when their subjective assessment of this delay makes it appear more favorable, by a potentially biased threshold. Note that this threshold, b' , is distinct from the proportional threshold defining a response window in the temporal general-

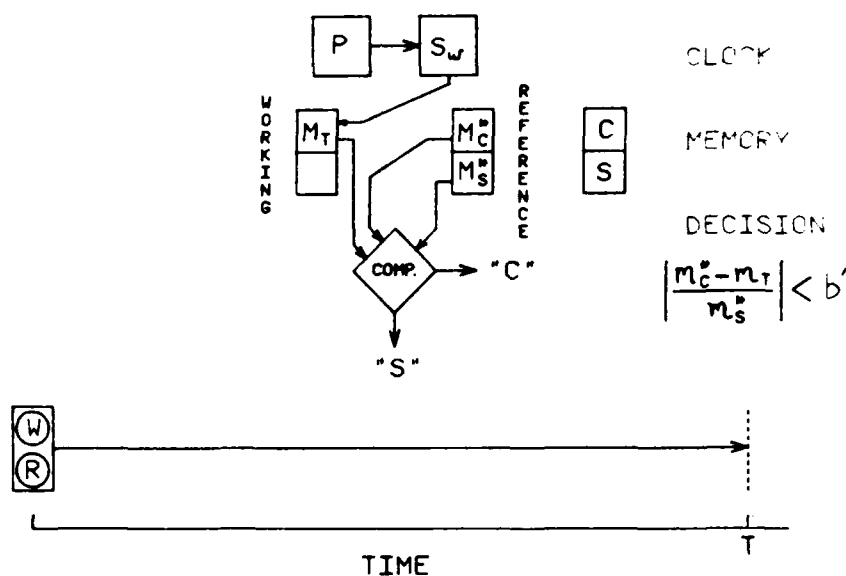


FIGURE 14. Information-processing model for choice in the "time left" task. During the choice period the switch gates pacemaker pulses to working memory for elapsed trial time, T . A comparison is made between the discrepancy between C and the current time—time left to food on the comparison side—and the standard delay to food, S . Preference is for the shorter of these two delays.

ization case. Here b' indicates a bias in favor of the time-left or standard side, with an unbiased mean value represented by $B' = 1.0$. We expect, however, that variance associated with the decision rule for both cases may be quite comparable.

The two terminal link conditions establish the memory values for the comparison and standard delays. In FIGURE 15, processing of the time-left terminal link is shown. The switch continues to gate pulses into the working memory for T , until reinforcement occurs when $T = C$ at the end of the trial. During this terminal link subjects are in a go/no-go situation precisely comparable to the left wing of a peak procedure (a fixed interval schedule). Hence their decision to respond or not to the white key is assumed to be based upon the time left to reinforcement normalized by the overall time to reinforcement, as shown in the decision rule in the lower right. After reinforcement, the reinforced value from working memory (m_C) is stored in reference memory (m_C^*) for the comparison interval.

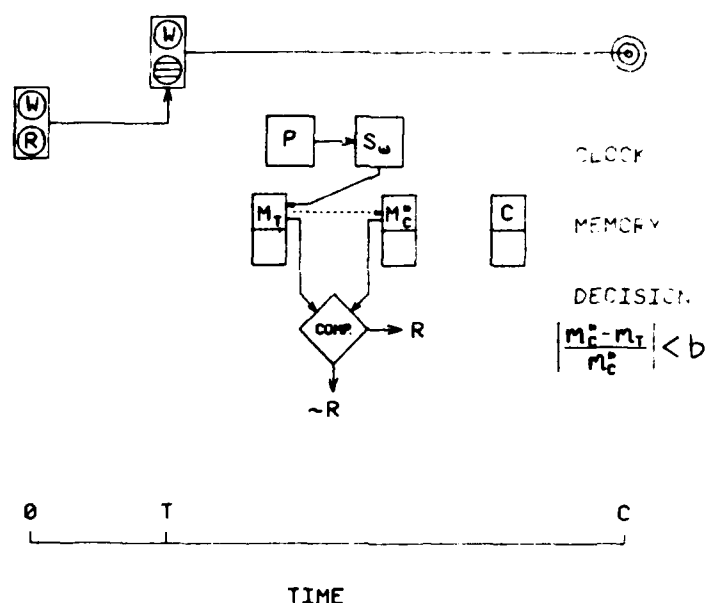


FIGURE 15. Information-processing model for the time-left terminal link. The switch continues to gate pulses to the trial-time accumulator, and the comparison is a go/no-go decision, as in the peak procedure.

When on other trials the choice at the entry point is in favor of the standard, the standard terminal link is entered. FIGURE 16 shows the switch gating pulses into a new working memory accumulator (m_i) as subjects begin timing the standard interval. Go/no-go responding is again assumed to be comparable to the left wing of a peak procedure, and the decision rule reflects the relative proximity to food. After reinforcement, the value in the working memory store (m_s) is transferred to a reference memory for the standard interval (m_s^*). In this way the memories for C and S are built

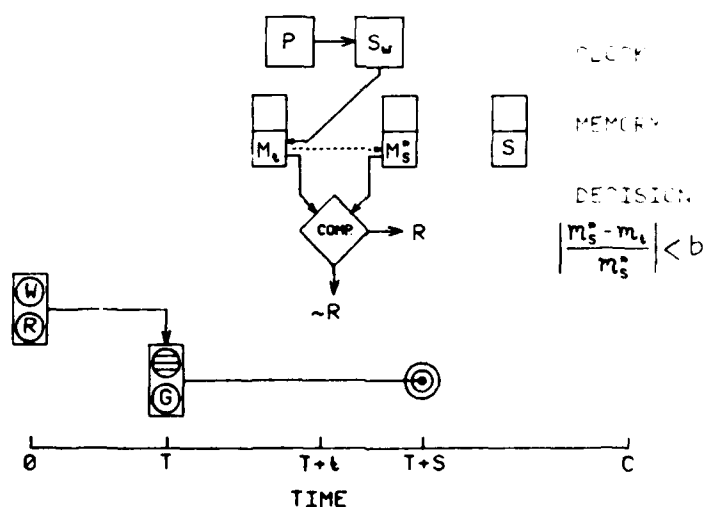


FIGURE 16. Information-processing model for the standard terminal link delay. During this delay the switch gates pulses into an accumulator for the standard delay in working memory, and the comparison is again a go/no-go decision, as in a peak procedure.

up over training and then become available for sampling during the choice period of later trials. In principle it should be possible to predict on individual trials the details of responding in the terminal links from features of responding in the choice link. We will pursue this line in later work. However, in what follows, we concentrate on analysis of preference, pooled over the choice period.

The time, $T_{1/2}$, at which subjects cross over during a sufficiently lengthy choice period from preferring S to preferring time-left is revealing. The decision rule implies that at this point the delays are subjectively equal. On average,

$$M_C^* - M_{T_{1/2}} = B^* M_S^*. \quad (2)$$

Translating these mean values according to Equation 1 gives us

$$T_{1/2} = \left(\frac{C}{S} - B^* \right) K^* S + T_0 [1 + K^* (B^* - 1)]. \quad (3)$$

Indifference points should be linear in S , for constant ratios of C to S , as described previously.¹⁵ Note that for $T_0 = 0$ and no bias ($B^* = 1$) or memory distortion ($K^* = 1$), $T_{1/2} = C - S$, when the actual time left, $C - T_{1/2}$, equals S . Deviations from physical equivalence may be introduced by bias and memory distortion. $T_{1/2}$ may be thought of as a set-point at which these factors have been subjectively equalized.

The performance of three subjects studied under three combinations of $S = C/2$ (7.5, 15; 15, 30; 30, 60) are shown in the three panels in FIGURE 17. The S value is

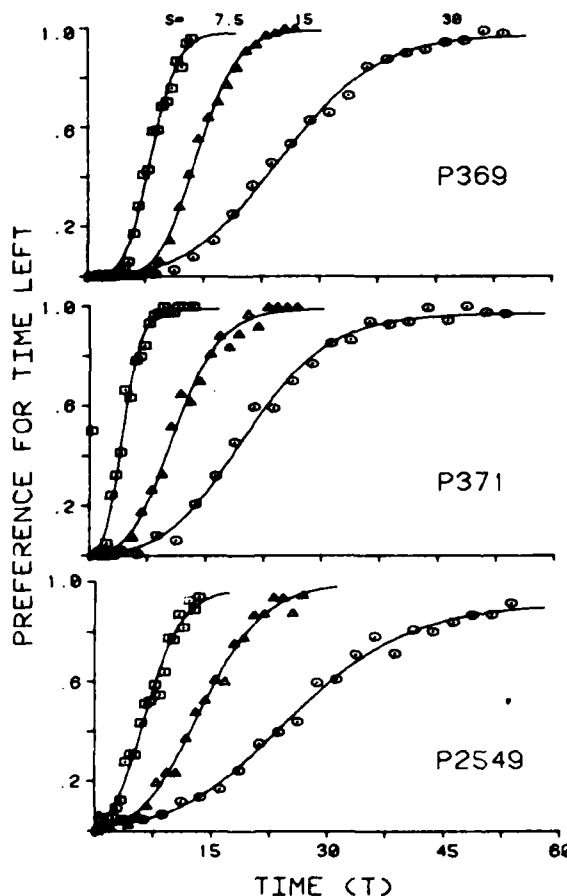
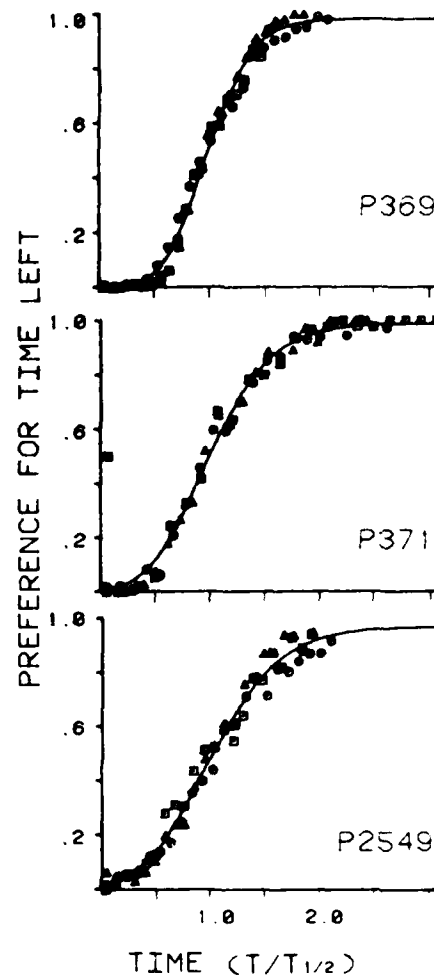


FIGURE 17. Psychometric preference functions for time left for three subjects at three different pairs of $S = C/2$ values. The functions are plotted against real time in the trial. The S parameter values are indicated in the top panel.

FIGURE 18. Preference functions for these three subjects replotted from FIGURE 17 as a function of time relative to the indifference point, $T_{1/2}$. The symbol code is the same as in FIGURE 17.



indicated above each curve in the top panel. These preference functions, like those published earlier,¹⁵ show a shallower rise as $S = C/2$ is increased, and $T_{1/2}$ values are approximately linear in S , as expected (Equation 3).

Our interest centers on a superposition property of the account, which is developed in the APPENDIX. There it is shown that if these curves are normalized by their $T_{1/2}$ values, the effects of changing bias ($B' \neq 1$) and memory encoding ($K^* \neq 1$) are cancelled in much the same way as the effects of changing K^* may be cancelled in the peak procedure through normalizing by K^*S . Here normalization by $T_{1/2}$ entails superposition as long as the scalar variance sources dominate. This will be true if the time ranges are large relative to T_0 and the Poisson rate parameter is greater than about 10 per sec. Plotting the preference functions at $T/T_{1/2}$ results in the superposition of these functions shown in FIGURE 18. Thus, again a Weber's law property is revealed, but with the important proviso that it operate on subjective, not objective time. Distortions in memory and innate preferences operate to rescale objective time.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Stephen Fairhurst and Amy Waring in all of the data and theory analysis presented here.

REFERENCES

1. GIBBON, J. & R. M. CHURCH. Sources of variance in an information processing theory of timing. Presented at the Harry Frank Guggenheim Conference on Animal Cognition, Columbia University, June 2-4, 1982. In press.
2. GIBBON, J. 1977. Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* **84**: 279-325.
3. CHURCH, R. M. & J. GIBBON. 1982. Temporal generalization. *J. Exp. Psychol. Animal Behav. Processes* **8**: 165-186.
4. HEINEMANN, E. G. 1982. A memory model for decision processes in pigeons. *In Quantitative Analyses of Behavior*. Vol. 3: Acquisition. M. Commons, R. J. Herrnstein, and A. R. Wagner, Eds. Ballinger. Cambridge, MA.
5. HEINEMANN, E. G. 1982. The presolution period and the detection of statistical associations. *In Quantitative Analyses of Behavior*. Vol. 4: Discrimination Processes. M. Commons, R. J. Herrnstein, and A. R. Wagner, Eds. Ballinger. Cambridge, MA.
6. TREISMAN, M. 1963. Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychol. Monogr.* **77**: 13.
7. CREELMAN, C. D. 1962. Human discrimination of auditory duration. *J. Acoust. Soc. Am.* **34**: 582-593.
8. CHURCH, R. M. 1984. Properties of the internal clock. This volume.
9. KRISTOFFERSON, A. B. 1967. Attention and psychophysical time. *Acta Psychol.* **27**: 93-100.
10. ALLAN, L. G., A. B. KRISTOFFERSON & E. W. WIENS. 1971. Duration discrimination of brief light flashes. *Percept. Psychophys.* **9**: 324-327.
11. WING, A. M. & A. B. KRISTOFFERSON. 1973. Response delays and the timing of discrete motor responses. *Percept. Psychophys.* **14**: 5-12.
12. GETTY, D. J. 1976. Counting processes in human timing. *Percept. Psychophys.* **20**: 191-197.
13. MCGILL, W. J. 1967. Neural counting mechanisms and energy detection in audition. *J. Math. Psychol.* **4**: 351-376.
14. GREEN, D. M. & R. D. LUCE. 1974. Counting and timing mechanisms in auditory discrimination and reaction time. *In Contemporary Developments in Mathematical Psychology*. Vol. 2: Measurement, Psychophysics, and Neural Information Processing. D. H. Krantz, R. C. Atkinson, R. D. Luce & P. Suppes, Eds. W. H. Freeman. San Francisco, CA.
15. ROBERTS, S. 1981. Isolation of an internal clock. *J. Exp. Psychol. Animal Behav. Processes* **7**: 242-268.
16. MECK, W. H. & R. M. CHURCH. 1981. Simultaneous temporal processing. *J. Exp. Psychol. Animal Behav. Processes* **7**: 242-268.
17. MECK, W. H. 1981. Selective adjustment of the speed of internal clock and memory processes. *J. Exp. Psychol. Animal Behav. Processes* **7**: 242-268.
18. GIBBON, J. & R. M. CHURCH. 1981. Time left: Linear versus logarithmic subjective time. *J. Exp. Psychol. Animal Behav. Processes* **7**: 87-108.

APPENDIX

Temporal Generalization and Peak

We require the total probability of the event in Equation 1, the decision rule, when all stages in the information-processing scheme may contribute variability. That is, we require the total probability of

$$\left| \frac{m_{S_i}^* - m_T}{m_{S_i}^*} \right| < b, \quad (A1)$$

where m_s^* , m_T , and b are each independent random variables sampled from reference memory, working memory, and from a threshold distribution on each trial. We develop this calculation by obtaining the mean and variance of the memory random variables, when these receive variance contributions from each of the postulated sources of variance in FIGURES 2 through 4. Then the analysis follows that of Church and Gibbon.³ At several points in the development normal approximations for distribution forms that are demonstrably slightly skewed will be used. We have in each case run computer simulations to satisfy ourselves that the approximations are reasonable. Our strategy is to add successive sources of variance moving through the information-processing chain from the pacemaker to the comparator.

The pacemaker generates pulses with Poisson variability at an intensity, λ . Imagine first that the effective switch closure time, d_T , is fixed at the mean value, D_T .^c

The number of counts accumulated in working memory associated with this switch closure time is then a Poisson-distributed variate with a mean and variance

$$E(m) = \text{Var}(m) = \lambda D.$$

We now allow the intensity parameter to vary. The moment generating function for the Poisson variate with fixed intensity is

$$M_{m_\lambda}(\theta) = \exp[\lambda D(e^\theta - 1)], \quad (\text{A2})$$

where we have replaced the subscript T by λ to indicate dependence on λ at any fixed T . If we now let λ vary around a mean Λ with standard deviation $\gamma\Lambda$, the moment generating function of the mixture then becomes

$$M_m(\theta) = \int_{-\infty}^{\infty} M_{m_\lambda}(\theta) f(\lambda) d\lambda. \quad (\text{A3})$$

Assuming normality of f , we may complete the square under the integral and factor so that λ is integrated out of (A3). With some rearrangement the moment generating function may be shown to be

$$M_m(\theta) = \exp [\Lambda D(e^\theta - 1) + 1/2 (\gamma \Lambda D)^2 (e^\theta - 1)^2]. \quad (\text{A4})$$

Taking first and second derivatives and setting $\theta = 0$ gives the first and second moments as

$$\begin{aligned} \mu'_1 &\equiv E(m) = \Lambda D, \\ \mu'_2 &= \Lambda D + (\Lambda D)^2 (1 + \gamma^2), \end{aligned} \quad (\text{A5a})$$

so that the variance of counts in working memory under both Poisson and scalar sources of variation is given by

$$\text{Var}(m) = \mu'_2 - \mu'^2_1 = \Lambda D + (\gamma \Lambda D)^2. \quad (\text{A5b})$$

The distribution form corresponding to the moment-generating function for the mixture (A4) is not normal, but approaches normality rapidly as T grows. In the ranges we will discuss, the forms are reasonably well approximated by normal distributions with mean and variance given by Equations A5a and 5b.

^cSubscripting will be dropped henceforth where the parameter dependence is obvious, and picked up again where needed. In the present case we assume the stimulus is on for an arbitrary but fixed period of time, T sec, and thus dependence on T will be suppressed.

We now wish to allow variability in the latency to open and close the switch. That is, we wish to obtain the mean and variance of the mixture that results from allowing d to vary normally about a mean, $D = T - T_0$ with standard deviation σ_0 . The conditional rules for mean and variance of x conditioned on y^d may be written

$$E(x) = E[E(x_y)], \quad (\text{A6a})$$

$$\text{Var}(x) = E[\text{Var}(x_y)] + \text{Var}[E(x_y)]. \quad (\text{A6b})$$

For the present case, allowing switch variance results in a mean and variance for working memory given by

$$E(m) \equiv M = \Lambda D, \quad (\text{A7a})$$

$$\text{Var}(m) \equiv \sigma_m^2 = a(\Lambda\sigma_0)^2 + \Lambda D + (a - 1)(\Lambda D)^2, \quad (\text{A7b})$$

where

$$a = 1 + \gamma^2. \quad (\text{A7c})$$

Working memory may be construed as involving a proportional translation for short-term storage of the count in the accumulator. Such a translation might add another source of variance. This source was considered in our earlier piece¹ and is not discriminable in form from variance induced by pacemaker rate variability, and hence we do not analyze it separately here.

Translation to reference memory occurs only on a subset of the trials, but after long training we assume the reference memory has built up a distribution of remembered, reinforced working memory values via the proportional translation indicated in FIGURE 4. The conditional rule (A6a) gives, for $m_k^* = k^* m_{S+}$,

$$E(m^*) = E[E(m_k^*)] \equiv M_{S+}^* = K^* \Lambda D_{S+}. \quad (\text{A8a})$$

And a lengthy but straightforward application of (A6b) gives, for $\sigma_{k+}^* = \gamma^* K^*$,

$$\text{Var}(m^*) = E[\text{Var}(m_k^*)] + \text{Var}[E(m_k^*)],$$

or

$$\sigma_{m+}^2 = K^{*2} [aa^*(\Lambda\sigma_0)^2 + a^*\Lambda D_{S+} + (aa^* - 1)(\Lambda D_{S+})^2], \quad (\text{A8b})$$

where

$$a^* = 1 + \gamma^{*2}. \quad (\text{A8c})$$

We are now in a position to analyze the decision rule, assuming independent, normally distributed memory variates.⁶

If we assume that the three variables in (A1) are independent and (essentially)

⁴Feller, W. 1966. An Introduction to Probability Theory and Its Applications. Vol. 1. 164 Wiley, New York, NY.

⁶The assumption of normality in the (product) memory variates violates somewhat the true character of these distributions, which have some positive skew. Again, however, in the range of seconds to minutes, this skew is not substantial and normal approximations are reasonable.

positive, the decision rule may be written

$$1 - b < \frac{m}{m^*} < 1 + b. \quad (\text{A9})$$

The correlated upper and lower limits (window edges) do not pose special difficulties here and may be treated as independent variables (compare the Appendix in Church and Gibbon³).

Defining $w_i = 1 + (-1)^i b$ with mean and variance,

$$E(w_i) \equiv W_i = 1 + (-1)^i B,$$

$$\text{Var}(w_i) = \sigma_b^2, \quad i = 1, 2,$$

Equation A9 becomes

$$m^* w_1 < m < m^* w_2. \quad (\text{A10})$$

Letting $x_i = m - m^* w_i$,

$$E(x_i) = M - M_S^* W_i. \quad (\text{A11a})$$

Repeated application of Equation A6b shows

$$\text{Var}(x_i) = \sigma_m^2 + \sigma_m^2 (\sigma_b^2 + W_i^2) + M_S^{*2} \sigma_b^2, \quad (\text{A11b})$$

where σ_m^2 and σ_m^* are given by Equations A7b and A8b. Assuming the x_i are normal, we have

$$P(R|T) = \Phi(Z_2) - \Phi(Z_1), \quad (\text{A12})$$

where

$$Z_i = \frac{-E(x_i)}{\sqrt{\text{Var}(x_i)}}, \quad i = 1, 2,$$

and Φ is the unit normal distribution function. It is convenient to divide numerator and denominator of the Z_i by M_S^* , giving

$$Z_i = \frac{W_i - \frac{D_T}{K^* D_S}}{\sqrt{\left(\frac{\sigma_m}{M_S^*}\right)^2 + \left(\frac{\sigma_m^*}{M_S^*}\right)^2 (\sigma_b^2 + W_i^2) + \sigma_b^2}}, \quad i = 1, 2. \quad (\text{A13})$$

Some algebraic rearrangement in Equations A7 and A8 gives

$$\left(\frac{\sigma_m}{M_S^*}\right)^2 = \frac{1}{(K^* D_S)^2} \left[a \sigma_0^2 + \frac{D_T}{\Lambda} + (a - 1) D_T^2 \right] \quad (\text{A14})$$

and

$$\left(\frac{\sigma_m^*}{M_S^*}\right)^2 = \frac{1}{D_S^2} \left[a^* a \sigma_0^2 + a^* \frac{D_S}{\Lambda} + (a^* a - 1) D_S^2 \right]. \quad (\text{A15})$$

where a^* and a are as defined in Equations A7 and A8. The denominator of Equation A13 with $i = 2$ was used for the window-edge standard deviation in FIGURE 9. The form (A12), with definitions (A13), (A14), and (A15), was used to plot the generalization gradients in the text. The peak procedure gradients require the additional scale factor $R_{\max}$, translating response probability into response rate,

$$\bar{R}_T = \bar{R}_{\max} P(R|T). \quad (\text{A16})$$

Time-Left Procedure

The time-left procedure is analyzed in a similar fashion, but with two memory variates, m_S^* and m_C^* , corresponding to reinforced exposure to the S and C delays. The memories are built up in the same way as in temporal generalization and (A8) and (A15) hold for $S+ = S, C$.

The decision rule for choice for time-left may be written:

$$m_C^* - m_T < b'm_S^*. \quad (\text{A17})$$

Defining

$$x = m_C^* - m_T - b'm_S^*,$$

$$P(R|T) = \Phi(Z), \quad (\text{A18})$$

where

$$Z = \frac{-E(x)}{\text{Var}(x)}, \quad (\text{A19a})$$

and

$$E(x) = M_C^* - M_T^* - b'M_S^*, \quad (\text{A19b})$$

$$\text{Var}(x) = \sigma_{m_C}^2 + \sigma_{m_T}^2 + \sigma_{m_S}^2 (\sigma_b^2 + b'^2) + M_S^{*2} \sigma_b^2. \quad (\text{A19c})$$

Again, it is convenient to divide numerator and denominator of (A19a) by M_S^* , giving

$$Z = \frac{b' - \frac{D_C}{D_S} + \frac{D_T}{K^* D_S}}{\sqrt{\left(\frac{\sigma_{m_C}}{M_S^*}\right)^2 + \left(\frac{\sigma_{m_T}}{M_S^*}\right)^2 + \left(\frac{\sigma_{m_S}}{M_S^*}\right)^2 (\sigma_b^2 + b'^2) + \sigma_b^2}}, \quad (\text{A20})$$

where $(\sigma_{m_i}/M_S^*)^2$ is defined by (A15) and

$$\left(\frac{\sigma_{m_i}}{M_S^*}\right)^2 = \left(\frac{1}{D_S}\right)^2 \left[a^* a \sigma_0^2 + a^* \frac{D_C}{\Lambda} + (a^* a - 1) D_C^2 \right]. \quad (\text{A21})$$

This form (A18, A20, and A21) was used in fitting the psychometric functions in figures in the text.

To see why near superposition is achieved when $P("C")$ is normalized by T , we

consider only large S , C , with S/C constant, so that T_0 may be neglected. From (A17) $T_{1/2}$ is given by

$$T_{1/2} = \left(\frac{C}{S} - B' \right) K^* S.$$

The numerator of (A20) becomes:

$$\left(\frac{C}{S} - B' \right) (T/T_{1/2} - 1),$$

which is constant at $T/T_{1/2}$.

The denominator terms may be analyzed separately. As S , C becomes large, the constant and Poisson components of variance become negligible and, from (A21)

$$\left(\frac{\sigma_{m_s}^*}{M_S^*} \right)^2 \rightarrow (a^* a - 1) \left(\frac{C}{S} \right)^2.$$

From (A15)

$$\left(\frac{\sigma_{m_s}^*}{M_S^*} \right)^2 \rightarrow (a^* a - 1),$$

and from (A14),

$$\left(\frac{\sigma_{m_T}}{M_S^*} \right)^2 \rightarrow (a - 1) \left(\frac{C}{S} - B' \right) (T/T_{1/2}).$$

Thus, the variance term approaches constancy for constant $T/T_{1/2}$ also, resulting in superposition in (A18).

Time Perception: Discussion Paper

R. DUNCAN LUCE

*Department of Psychology and Social Relations
Harvard University
Cambridge, Massachusetts 02138*

Study of the papers in this session reveals at least three major topics, each of which arises in three or more of the papers. They are: the quality of timing performance, clock models aimed at accounting for the variability in the behavior, and discovery of the scale for the subjective perception of time. I organize my remarks accordingly.

QUALITY OF TIMING PERFORMANCE

Perhaps the most obvious contrast in quality of timing is that between human and animal performance. In both cases considerable evidence is provided that mean times are quite accurate and that, to a first approximation anyhow, all distributions of normalized responses are the same. However, substantial differences exist between the animal and human data in the magnitude of the relative variability: the Weber fraction for the human data runs at about 5% and for the animals nearer to 50%.¹⁻³ However, in at least two respects these two classes of data are not comparable. First, the ranges over which they have been studied do not overlap, being between tens of milliseconds and a few seconds for the humans and from seconds to tens of seconds for the animals. Second, the pressure in the human experiments has been for precision of performance and it is far from clear that the animal studies have been designed with that in mind. The consequences for an animal who does not exhibit exact timing are really not very severe, being nothing worse than some unrewarded responding. Perhaps it would be useful for someone doing animal studies to attempt to establish the limits of their performance, which we have no reason to expect to be worse by an order of magnitude than that of people. And in the other camp, perhaps it would be useful to determine whether the 5% figure continues to hold into the region of tens of seconds. I do not underestimate the difficulties and effort required in each case, but both questions seem important.

All of us have been astonished by the precision of timing that Kristofferson and his associates have managed to achieve, and until he demonstrated it 15 years ago few of us would have anticipated that the variability in timing would remain constant or nearly so over any substantial region. What is new in the present data, and even more surprising, is the series of plateaus in estimates of the period of the clock (see below), which are spaced by factors of two over intervals that increase by factors of two. This, it seems to me, has important implications for modeling, to which I now turn.

CLOCK MODELS

A number of our authors envisage timing behavior as based upon a clock of some sort. Three of the papers^{1,2,4} postulate a real-time digital device with timing arising from a count of the number of events. Hopkins and Kristofferson admit no variability in the clock itself, whereas Gibbon *et al.* explore various possibilities, rejecting as a primary source of the observed variability Poisson noise in the clock and favoring some form of scalar variability either in the clock or in the memory. (I should make clear that their

finding that Poisson variability plays little role in timing in no way bears on the physiologically well justified Poisson representation of sensory intensity.) Assuming the period of the clock to be q , everyone agrees that the arrival of a signal will be random relative to the pulse train defining the clock, which introduces a uniformly distributed random variable over the interval $(0, q)$. Wing and Kristofferson⁵ suggested that this is just one of three sources of variability leading to the observed variability, the other two being another, but independent, uniformly distributed one also on $(0, q)$, and the third an independent, normally distributed one associated with the response process. Hopkins has shown us that this model gives an almost perfect fit to his data; however, one would also like to see how well the data can be fit by other highly peaked distributions, such as the Laplace, instead of the triangular.

The major problem of that model, it seems to me, is this: Where does the second uniform distribution over $(0, q)$ come from? Hopkins attempted an argument along the following lines: After the count is achieved, the system exits the clock and initiates a response mechanism which is delayed in starting in much the same way the clock is, presumably because it cycles in a clock-like fashion. To fit the data, the two rates must be nearly the same, but to achieve approximate independence he assumed slightly different rates. This argument does not seem very persuasive to me, and I fear that it may run into difficulties with Kristofferson's findings about the plateaus.

Consider how the plateaus of variance may come about. One possibility is that the counter applied to the pulses of the clock has a maximum count, and when a time is wanted that exceeds the capacity of the counter, the system in essence counts every other pulse. This could be achieved by cells that are activated whenever two pulses occur within q time units but are refractory for considerably less than q time units, where we recall q is estimated to be about 12 msec. Such a model produces one uniformly distributed random variable on the interval $(0, 2q)$, but I really don't see where the second one is to come from since there is no reason for the quantal character of the response process also to change scale. Because the second random variable seems to arise from exiting the clock and initiating the response process, its distribution should be controlled by the statistics of the response mechanism, not that of the timer. Once that dilemma is solved, then estimating even longer intervals simply involves repeated applications of the same type of cell that responds to every other pulse, but with even broader periods of integration. Such a mechanism generates the factors of two which Kristofferson has found. One cannot but wonder how many of these filtering cells can be arranged in series; presumably that can be estimated by extending Kristofferson's methods to appreciably longer times. It seems important to me that the distributions for $2q$ and $3q$ be studied with the same care the Hopkins has given q to see whether the fit of the convolution of a normal with two identical, independent uniform distributions continues to be equally satisfactory. For the reasons given above, I wonder if an asymmetry will not begin to evidence itself. If it does not, the second uniform distribution on $(0, 2q)$ is an interesting theoretical challenge.

Before I turn to my last topic, let me say how pleased I am to find growing evidence for the existence of both good mental clocks and accurate mental counters, which some years ago David Green and I⁶ suggested would provide a parsimonious account of some psychophysical speed-accuracy data.

SUBJECTIVE PERCEPTION OF TIME

When we turn to the subjective aspects of the perception of time, the only phrase that comes to my mind is "a can of worms." It is a familiar can to those who, like myself, have fished in psychophysical waters.

With the exception of Eisler (see below), those who have spoken of clock models have postulated periodicity in real time, and to the degree that the models are successful, which is considerable, that can be taken as *prima facie* evidence that at a certain level the perception of time is proportional to physical time. In this view, subjective scales are no more than useful constructs in a theory, and certainly many important constructs of physics—energy, momentum, entropy, and force—gained their status only via theory. However, as psychologists we have, in addition, strong intuitions about the lively existence of subjective attributes that cannot possibly be linear with the usual physical measures as well as the added knowledge that when we ask human subjects, more or less directly, about these attributes, we usually obtain results that are far from linear with physical measures. That makes suspect, but by no means rules out, the proportionality of subjective to physical time which is posited, with success, in these models.

Some⁵ observe that the distributional data are describable as arising from a single distribution through scale changes, which is what Weber's law amounts to, and suggest that this in essence determines the needed transformation of time—which transformation is located in memory and not in the clock. This is the original strategy of Fechner, one that postulates a solution which, at least in psychophysics, has been found wanting an empirical basis.

So, one says, almost reflexively, let's decide the matter empirically. It is perhaps well to begin with the blunt admission that psychophysicists have never evolved a way to do so that has commanded wide assent. The direct scaling methods of S.S. Stevens⁷ to which Eisler made reference, rest upon a mode of communication that is entirely language based; in fact, these methods rely upon the instruction to the subjects that the numbers they assign to stimuli shall preserve subjective ratios. In whatever way our subjects understand this instruction, they do give consistent, repeatable data; nonetheless, whatever the instruction does mean, we do not have the slightest idea how to communicate it nonverbally. Moreover, through the work of King and Lockhead,⁸ we know that magnitude estimates are highly malleable, and quite different functions can be obtained by altering the feedback subjects receive. In brief, we simply do not know how to do scaling experiments with subjects who do not speak our language. Yet, that is exactly what two, and perhaps three, of our authors have claimed to be doing.^{1,3} Do I misunderstand and have they solved the century-old dilemma of the psychophysics of big differences, of what I call global psychophysics? I think not.

So far as I can tell, the researchers working with animals are doing temporal discrimination studies which, just like the discrimination studies of psychophysics, do not tell us much about the overall apprehension of an attribute. The fact that the indifference point between a variable and a standard time interval sometimes is approximately at the standard in no way implies that a linear scale is involved, and the fact that Weber's law holds does not dictate a particular nonlinear transformation. Eisler is quite aware that neither tack will do, but I believe he has slipped into two other traps. First, he has avoided reducing the problem to one of simple discrimination by assuming that the subject selects the second interval not to be equal to the first one, but to be subjectively one-half of the total interval. The motivation for this bit of indirection on the part of the animal, although not the author, escapes me. How does the animal know to use 1/2 rather than any other fraction? Second, and rather more serious, he has used the human data to establish a region within which he reinforced the animals' responses in what amounts to a discrimination study with 10 discriminative stimuli, and the animals—at least two of eight rats—quite reasonably took into account their own variability and stayed well within the reinforced region, thereby nearly reproducing the human behavior. Since we know from years of operant work that animals are quite sensitive to temporal reinforcement and from human work that

magnitude estimation scales are malleable, these results persuade me of nothing whatsoever about temporal perception in animals.

I do not wish to disparage efforts toward finding objective ways of eliciting information about internal states, which is what I believe a subjective scale to be, but it is surely going to require a more complex idea than either just discrimination or just reinforcement. For the moment we may have to be satisfied with models of the sort that Kristofferson and his students and Gibbon and Church have been working on to account for these highly regular and, I believe, important temporal discrimination and timing data.

REFERENCES

1. GIBBON, J., CHURCH, R. M. & W. H. MECK. 1984. Clock, memory, and decision processes in different timing. This volume.
2. KRISTOFFERSON, A. B. 1984. Quantal and deterministic timing in human duration discrimination. This volume.
3. STUBBS, D. A., L. R. DREYFUS & J. G. FETTERMAN. 1984. The perception of temporal events. This volume.
4. HOPKINS, G. W. 1984. Ultrastable stimulus-response latencies: Towards a model of response-stimulus synchronization. This volume.
5. WING, A. M. & A. B. KRISTOFFERSON. 1973. Percept. Psychophys. 14: 5.
6. GREEN, D. M. & R. D. LUCE. 1973. In Attention and Performance: IV. S. Kornblum, Ed.: 547-569. Academic Press. New York, NY.
7. STEVENS, S. S. 1975. Psychophysics. Wiley. New York, NY.
8. KING, M. C. & G. R. LOCKHEAD. 1981. Response scales and sequential effects in judgment. Percept. Psychophys. 30: 559-603.
9. EISLER, H. 1984. Subjective duration in rats: The psychophysical function. This volume.

The Representation of Mental Activities in Critical Path Networks

RICHARD J. SCHWEICKERT

*Department of Psychological Sciences
Purdue University
West Lafayette, Indiana 47907*

Suppose a subject executes a number of mental processes to perform a task. We cannot observe the processes directly, but we can observe the time required to complete the task under various conditions. The problem is to determine, as far as possible, how the processes are organized and how much time each process takes. I will summarize an approach to this problem based on the theory of scheduling and on Sternberg's¹ additive factor method. Then this approach will be used to locate the decision process in several cognitive tasks.

The theory of scheduling deals with the optimal organization of processes in computers, factories, and so on.²⁻⁵ One common type of process organization is the critical path network. Each process in the task is represented by an arrow in the network, and if process *a* must be completed before process *b* can start, then the arrow corresponding to *a* precedes the arrow corresponding to *b* (Fig. 1). No arrow is allowed to precede itself, and the network is said to be acyclic.

Processes in a critical path network are related in one of two ways. Those joined by a directed path, such as *a* and *b* in FIGURE 1, must be executed in order and are called *sequential*. Those not joined by a directed path, such as *a* and *c* in FIGURE 1, are called *concurrent* because they can be executed simultaneously. Note that processes in series and in parallel⁶ are special cases of sequential and concurrent processes, respectively. Two sequential processes are in *series* if a process is on a path with one of them if and only if it is on a path with the other. Two processes are in *parallel* if they have the same starting and terminating points. For example, *a* and *c* are in parallel in FIGURE 1. We assume that no process can begin until all those preceding it are finished. Every process has a *duration*, and the duration of the task, the response time, is the sum of the durations of the processes on the longest path through the network, called the *critical path*.

As the critical path method is ordinarily used, the network is given, and the time required to complete the task is to be calculated. We have the opposite problem: We know the time required to complete the task under various conditions and want to construct the unknown network. The key to constructing the network is to use the idea from Sternberg's additive factor method of prolonging processes. The effects of such prolongations are surprisingly informative about the network.

The assumptions here differ from those of the additive factor method in that we allow for the possibility of two processes being executed concurrently. The assumptions are also different from those of McClelland's⁷ cascade model. In his model, more than one process is in execution at a time, but the processes are sequential in the sense that information is passed continuously, from one process to the next, in order.

A further assumption will be made: that the process durations are fixed quantities and do not vary from trial to trial. Equations based on this assumption are only approximations to the correct stochastic equations, but with long prolongations and a large number of trials the approximations are not too bad.⁸ Without this assumption, the problems are formidable, although some progress has been made.⁹⁻¹¹

With Sternberg's¹ additive factor method, two experimental factors which increase reaction time are manipulated. Suppose that (a) all the processes of a task are executed in a sequence and that (b) each factor prolongs a different process. Then the effect on reaction time of prolonging both processes will be the sum of the effects of prolonging them individually. A violation of additivity has usually been interpreted in the framework of the method as indicating that (b) is false. But nonadditivity might indicate instead that (a) is false, and that processing is not entirely sequential. Nonadditivity turns out to be likely when separate processes are prolonged in a task involving concurrent processing. For such a task, of course, there is no reason to expect the principles of the additive factor method to apply without modification.

Latent Network Theory

The increase in response time produced by prolonging two processes depends on whether they are concurrent or sequential. Let T denote the response time when all the processes are at the shortest durations used in the experiment. Let $\Delta T(\Delta x, 0)$ denote the increase in T produced by prolonging process x by Δx , leaving y unchanged; other

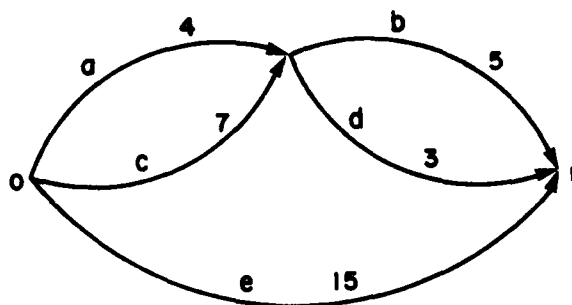


FIGURE 1. Each arrow represents a mental process that must be executed to perform a task. The numbers on the arrows are the process durations.

increases are denoted analogously. It can be shown that if x and y are concurrent, then

$$\Delta T(\Delta x, \Delta y) = \max [\Delta T(\Delta x, 0), \Delta T(0, \Delta y)]. \quad (1)$$

(All the equations in this section were derived in an earlier paper.¹²)

The situation is more complicated if x and y are sequential. Suppose x precedes y . The amount of time by which x can be prolonged without making y start late is called the *slack* from x to y , written $s(xy)$. Similarly, the amount of time by which x can be prolonged without delaying the response, r , and thereby increasing the response time is called the *total slack* for x , written $s(xr)$. A process is on a *critical path* if and only if its total slack is zero. If all the processes are in a sequence, there is only one path, necessarily critical, so every process has zero total slack.

Slack is important when two sequential processes are prolonged. Suppose x precedes y on a path. If the prolongations Δx and Δy are not too small, then it can be shown that

$$\Delta T(\Delta x, \Delta y) = \Delta T(\Delta x, 0) + \Delta T(0, \Delta y) + k(xy), \quad (2)$$

where $k(xy) = s(xr) - s(xy)$ is called the *coupled slack* from x to y .

The magnitude of $k(xy)$ does not depend on the magnitudes of Δx and Δy . This fact provides a strong test of whether a network analysis applies to a given set of data: All values of Δx and Δy large enough for Equation 2 to hold should yield the same value for $k(xy)$, the interaction term.

If all the processes are sequential, $k(xy) = 0$ for every pair x and y , and Equation 2 becomes the additive relationship of the additive factor method. In general, however, Equations 1 and 2 indicate that when two separate processes in a network are prolonged, their effects will interact.

The Wheatstone Bridge

A negative value of $k(xy)$ is very informative. If x precedes y and $k(xy) < 0$, then the task network must have a subnetwork in the shape illustrated in FIGURE 2, called a Wheatstone bridge. Moreover, certain relationships hold among the path durations, although these are not relevant here. See Schweickert¹² for the details and proof.

A peculiarity of processes x and y arranged in a Wheatstone bridge with $k(xy) < 0$ is that small prolongations of x and y will result in the holding of Equation 1 rather than 2. That is, for small prolongations, x and y will mimic concurrent processes. This can only occur with a Wheatstone bridge. This mimicking of concurrent processes by

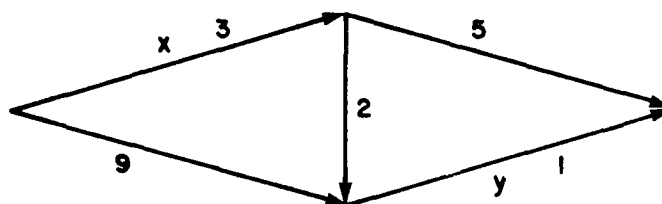


FIGURE 2. A Wheatstone bridge. If the coupled slack for x and y is negative, then the task network has a subnetwork of this shape.

sequential ones is the analog of the nonidentifiability discovered by Townsend,⁶ that serial and parallel processes often cannot be distinguished on the basis of their completion time distributions.

Determining Processing Order

Experiments in which the subject makes two responses on every trial are also informative. Let the two response times be T_1 and T_2 , both measured from the same point. Each response time considered alone will satisfy Equation 1 or 2 under the appropriate conditions. Furthermore, T_1 and T_2 are related. Suppose x precedes y , which precedes both responses, and the prolongations Δx and Δy are not too small. Then

$$\Delta T_1(\Delta x, \Delta y) - \Delta T_1(0, \Delta y) = \Delta T_2(\Delta x, \Delta y) - \Delta T_2(0, \Delta y). \quad (3)$$

If, instead, y precedes x , then $\Delta T_1(\Delta x, 0)$ and $\Delta T_2(\Delta x, 0)$ are required in Equation 3 in place of $\Delta T_1(0, \Delta y)$ and $\Delta T_2(0, \Delta y)$. The order of x and y is revealed, then, if one version of Equation 3 holds but not the other. If neither version holds, a network model is invalid.

There is another way to find the order of execution of processes. Suppose x precedes y which precedes z . If the prolongations Δx , Δy , and Δz are not too small, then the combined effect of prolonging all three processes is

$$\Delta T(\Delta x, \Delta y, \Delta z) = \Delta T(\Delta x, 0, 0) + k(xy) + \Delta T(0, \Delta y, 0) + k(yz) + \Delta T(0, 0, \Delta z). \quad (4)$$

This equation is useful for two reasons. First, since all the parameters in it can be determined by prolonging the processes individually and in pairs, the equation makes a prediction which can be tested. Second, the equation gives information about order: y is executed between x and z . To see this, note that for the three processes there are three coupled slacks, $k(xy)$, $k(yz)$ and $k(xz)$, but $k(xz)$, the one corresponding to the first and last of the three processes, is missing in the above equation. If the order were, say x , then z , then y , the missing term would be $k(xy)$.

The method summarized here has three major advantages over most other methods for analyzing reaction times: (1) It can be used even if the processes are not all in series. Furthermore, one can usually determine whether two mental processes are sequential or concurrent. (2) If processes are sequential, one can often determine their order of execution. (3) Information about the durations of processes is provided by the magnitudes of the coupled slacks, $k(xy)$ (see below).

APPLICATION: THE LOCATION OF DECISIONS IN COGNITIVE TASKS

The procedure just described will be applied to four information-processing tasks, each of which involves a decision, that is, a process prolonged by increasing the information in the stimulus.^{13,14}

Digit-Naming

Two digit naming experiments will be discussed, one by Sternberg²⁰ (p. 296) and a closely related one by Blackman.¹⁵ In Sternberg's experiment, subjects were presented visually with a digit and responded with a spoken digit. Three factors were manipulated: (i) Stimulus quality was degraded by superimposing a checkerboard pattern over the digit on some trials. Let E be the process prolonged by this procedure. (ii) The number of alternative digits was two in some blocks of trials and eight in others. We will call the process prolonged by increasing the number of alternatives the decision, D . (iii) On some trials the subject named the digit, and on other trials he named the digit plus one. Let Q be the process prolonged when the subject adds one to the digit.

The data are in TABLE 1. The effects of degrading the stimulus and of adding one are additive, indicating that each of those manipulations affects a separate process; the two processes are sequential. The effects of degrading the stimulus and increasing the number of alternatives interact, as do the effects of increasing the number of alternatives and adding one to the digit.

The additive factor explanation of these results assumes that the task is composed of a sequence of processes and degrading the stimulus prolongs one of them, E , while adding one to the digit prolongs another, Q . If increasing the number of alternatives prolonged a third process in the sequence, then the effects of this factor should be additive with the others. The explanation, according to the additive factor method, of the interactions which were found instead is that increasing the number of alternatives

TABLE 1. Changes in Response Times in Sternberg's Digit-Naming Experiment^a

Factor Levels				
Stimulus Quality	Number of Alternatives	Transformation	Prolongations	ΔT
Degraded	8	$x + 1$	$\Delta E \Delta D \Delta Q$	197
Intact	8	$x + 1$	$\Delta D \Delta Q$	144
Degraded	2	$x + 1$	$\Delta E \Delta Q$	45
Intact	2	$x + 1$	ΔQ	18
Degraded	8	x	$\Delta E \Delta D$	97
Intact	8	x	ΔD	43
Degraded	2	x	ΔE	30
Intact	2	x	Baseline	0

NOTE: The baseline reaction time was 328 msec.

^aData are from Figure 8 of Sternberg.¹

affects both E and Q . If we relax the assumption that all the processes are in a sequence, however, the interactions can be explained if each factor affects a process in a critical path network.

Suppose the processes are arranged in a network. The data indicate that Equation 2 holds for every pair of the processes E , D , and Q . The values of the coupled slacks are calculated from TABLE 1.

$$k(EQ) = \Delta T(\Delta E, 0, \Delta Q) - \Delta T(\Delta E, 0, 0) - \Delta T(0, 0, \Delta Q) = -3,$$

$$k(ED) = \Delta T(\Delta E, \Delta D, 0) - \Delta T(\Delta E, 0, 0) - \Delta T(0, \Delta D, 0) = 24,$$

$$k(DQ) = \Delta T(0, \Delta D, \Delta Q) - \Delta T(0, \Delta D, 0) - \Delta T(0, 0, \Delta Q) = 83.$$

These equations indicate that processes E , D , and Q are all on a path together. What is their order? Common sense suggests that E , the process affected by stimulus degradation, would come first. In his discussion of the experiment, Sternberg says that degrading the stimulus and having the subject add one to the digit probably affect processes that are widely separated. This suggests that D is in the middle, and the order is E , then D , and then Q . This order can be tested by Equation 4, which should hold in the following form if the order E, D, Q is correct:

$$\begin{aligned} \Delta T(\Delta E, \Delta D, \Delta Q) = & \Delta T(\Delta E, 0, 0) + k(ED) \\ & + \Delta T(0, \Delta D, 0) + k(DQ) + \Delta T(0, 0, \Delta Q). \end{aligned}$$

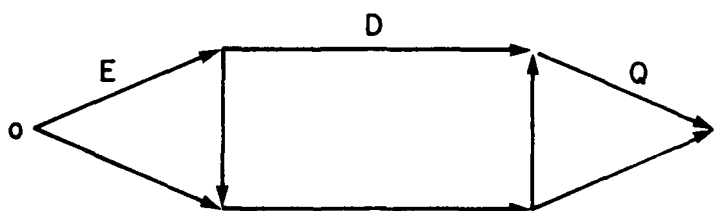


FIGURE 3. A network for the digit-naming experiments of Sternberg and Blackman. The onset of the digit is at o and the response is made at r . Process E is prolonged by stimulus degradation, D is the decision, and Q is prolonged when the subject adds one to the digit.

From TABLE 1,

$$197 \approx 30 + 24 + 43 + 83 + 18.$$

Since the right hand side is 198, the error is 1 and the equation holds.

An equally good fit would be given by the order Q, D, E , although this order is unintuitive. Every other order does worse. The fact that Equation 4 holds for some order is, of course, support for a network model. A network representing the mental processes in this task is given in FIGURE 3. Details about its construction and about the process durations are in the APPENDIX.

A Related Digit-Naming Task by Blackman

Blackman¹⁵ performed an experiment which is a replication of the one by Sternberg, with one change. Sternberg increased the number of alternative digits from

TABLE 2. Changes in Response Times in Blackman's Digit-Naming Experiment¹⁵

Factor Levels				
Stimulus Quality	Stimulus Probability	Transformation	Prolongations	ΔT
Degraded	.15	$x + 1$	$\Delta E \Delta D \Delta Q$	307
Intact	.15	$x + 1$	$\Delta D \Delta Q$	146
Degraded	.70	$x + 1$	$\Delta E \Delta Q$	137
Intact	.70	$x + 1$	ΔQ	31
Degraded	.15	x	$\Delta E \Delta D$	227
Intact	.15	x	ΔD	76
Degraded	.70	x	ΔE	97
Intact	.70	x	Baseline	0

NOTE: The baseline reaction time was 445 msec.

2 to 8 and thereby prolonged a process we called D . Blackman always presented three stimuli, but varied their probabilities. The most probable stimulus was presented 70% of the time and each of the other two was presented 15% of the time.

We will suppose that the process affected by stimulus probability is the same process as that affected by the number of alternatives, since manipulating the latter is a special case of manipulating the former, and we will call this process the decision, D .

The data for Blackman's experiment are given in TABLE 2. The coupled slacks have the following values:

$$k(EQ) = \Delta T(\Delta E, 0, \Delta Q) - \Delta T(\Delta E, 0, 0) - \Delta T(0, 0, \Delta Q) = 9,$$

$$k(ED) = \Delta T(\Delta E, \Delta D, 0) - \Delta T(\Delta E, 0, 0) - \Delta T(0, \Delta D, 0) = 54,$$

$$k(DQ) = \Delta T(0, \Delta D, \Delta Q) - \Delta T(0, \Delta D, 0) - \Delta T(0, 0, \Delta Q) = 39.$$

Since Equation 2 holds for each pair of processes, E, D , and Q are all on a path together. To be consistent with the results found in Sternberg's experiment, their order should be E , then D , then Q . Equation 4 should then hold in the form

$$\Delta T(\Delta E, \Delta D, \Delta Q) = \Delta T(\Delta E, 0, 0) + k(ED) \\ + \Delta T(0, \Delta D, 0) + k(DQ) + \Delta T(0, 0, \Delta Q).$$

From TABLE 2,

$$307 \approx 97 + 54 + 76 + 39 + 31,$$

the error is 10, and the equation holds.

Since the experiments of Sternberg and Blackman are so similar, one would expect that the networks representing them would be similar, and in fact, the network in FIGURE 3 represents both; the details are in the APPENDIX.

A Memory-Scanning Task

The experiments by Sternberg and Blackman support, through Equation 4, the idea that stimulus quality, stimulus probability and stimulus-response compatibility

TABLE 3. Response Times in the Experiment of Miller and Anbar¹⁶

Degradation	Stimulus Probability		
	Baseline .31	$\Delta_1 D$.15	$\Delta_2 D$.04
	<i>Unexpected</i>		
$\Delta_2 E$ (dim)	849	894	925
$\Delta_1 E$ (dots)	709	786	859
Baseline (normal)	648	659	731
	<i>Expected</i>		
$\Delta_2 E$ (dim)	757	780	823
$\Delta_1 E$ (dots)	622	643	685
Baseline (normal)	547	551	585

NOTE: Data published in Figures 4 and 5 of Miller and Anbar.¹⁶ Numerical values from Miller, J. Personal communication, May 14, 1982.

affect processes arranged in a critical path network. Further support, this time through Equation 2, comes from an experiment by Miller and Anbar¹⁶ (experiment II).

The task was memory scanning using letters as stimuli. Three letters were in the positive set. On each trial, the subject was presented with a probe letter and indicated, by pressing a button, whether the probe was an element of the positive set or not. Stimulus probability and stimulus quality were manipulated. As above, we denote the process affected by stimulus quality as *E* and that affected by stimulus probability as *D*. A third factor, expectancy, was also manipulated.

The effects of decreasing expectancy do not agree with the equations we would expect to hold if it were prolonging a single process in a critical path network, although a critical path network describes the effects of the other two factors. The effects of expectancy were somewhat irregular, and I will not discuss them further, although the data are given in TABLE 3 for the interested reader. The data discussed here are for the condition in which the stimuli were expected.

Three levels of stimulus probability were used, .31, .15 and .04. Stimulus encoding was affected in two different ways, by decreasing contrast with a filter over the screen

AD-A150 031

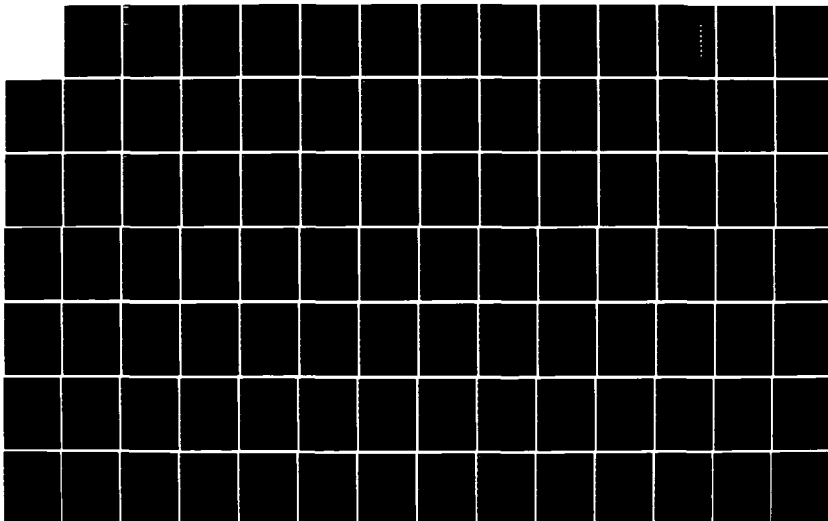
ANNALS OF THE NEW YORK ACADEMY OF SCIENCES VOLUME 423
TIMING AND TIME PER. (U) NEW YORK ACADEMY OF SCIENCES
NY J GIBBON ET AL. 13 MAY 84 N00014-84-G-0130

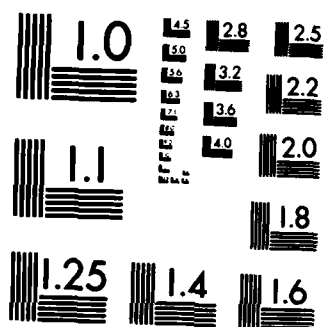
2/3

UNCLASSIFIED

F/G 5/10

NL





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

and by superimposing dots over the stimuli. The data are consistent with the idea that changing the contrast and superimposing the dots are two levels of a factor that prolongs a single process, *E*. It will be convenient for us to accept this idea, although an analysis of the results carried out without this assumption would also support a network model.

The data are in TABLE 3. Note that when the letters are presented normally, there is little effect of changing stimulus probability from .31 to .15, but there is an effect of changing stimulus probability from .31 to .04. The network model explanation is that *D*, the process prolonged by decreasing stimulus probability, has slack. A small prolongation of *D* has no effect on response time, while a larger prolongation has an effect.

Let $\Delta_1 D$ be the amount by which *D* is prolonged when stimulus probability is decreased from .31 to .15, and let $\Delta_2 D$ be the amount by which *D* is prolonged when stimulus probability is decreased from .31 to .04. The combined effects of changing stimulus quality and stimulus probability are described by Equation 2. That is,

$$T(\Delta_1 E, \Delta_2 D) - T(\Delta_1 E, 0) - T(0, \Delta_2 D) + T(0, 0) = k(ED)$$

$$685 \quad - \quad 622 \quad - \quad 585 \quad + \quad 547 \quad = \quad 25.$$

Furthermore, in accordance with the model, the larger level of prolongation of *E*, $\Delta_2 E$, leads to the same value for $k(ED)$,

$$T(\Delta_2 E, \Delta_2 D) - T(\Delta_2 E, 0) - T(0, \Delta_2 D) + T(0, 0) = k(ED)$$

$$823 \quad - \quad 757 \quad - \quad 585 \quad + \quad 547 \quad = \quad 28.$$

Since the two values of $k(ED)$ are about the same ($25 \approx 28$), the equations above support the idea that *E* and *D* are sequential processes in a critical path network.

At the normal level of stimulus quality, changing stimulus probability from .31 to .15 had little effect on response time. But TABLE 3 shows that beyond the normal level, the effects of changing stimulus quality and of decreasing stimulus probability have additive effects on response time. The additivity is found for decreasing stimulus probability from .31 to .15, and also for decreasing it from .31 to .04. In other words, the coupled slack between *E* and *D* is 0, and in accordance with Equation 2 this value does not depend on the level of prolongation of *D*. It is likely that at these levels for the duration of *E* and *D*, *E* and *D* are critical processes. The coupled slack is zero because factors prolonging critical processes are additive factors.

A Dual Task

I will now discuss the location of a decision in a more complex task.¹⁷ In a dual tone and digit identification task, Becker¹⁸ investigated the joint effects of the size of the interstimulus interval and the difficulty of the tone decision. The digit 1 or 2 was presented visually, and after an interstimulus interval of either 90 or 190 msec a high- or low-frequency tone was presented. With his left hand the subject pressed one of two buttons to indicate which digit occurred. With his right hand he responded to the tone. The decision about the tone was manipulated by requiring either one or two alternative responses to the tone. In the one alternative condition, the subject pressed the same button when either tone occurred, and in the two alternative condition he pressed one of two buttons to indicate which tone occurred.

Let *I* be the interstimulus interval and let *N* be the process prolonged by increasing the number of alternatives for the tone response. We will call *N* the decision about the

tone. When there are two responses, it is customary to measure the reaction time for the second response from the onset of the second stimulus. For our purposes, it is convenient to measure all the times with respect to the same point, namely the onset of the first stimulus, the digit. Let o_d denote this point. Let $T_d(0, 0) = 346$ be the baseline response time to respond to the digit and let $T_n(0, 0) = 475$ be the baseline response time to respond to the tone, both time intervals measured with respect to o_d .

The time to respond to the digit when I and N are prolonged will be denoted $T_d(\Delta I, \Delta N)$. The increase in response time to the digit when I and N are prolonged will be denoted $\Delta T_d(\Delta I, \Delta N) = T_d(\Delta I, \Delta N) - T_d(0, 0)$, and so on. (In each condition T_d is the same as RT1 in Table 1 of Becker's article, and T_n is RT2 plus the appropriate ISI.) TABLE 4 gives the response times and changes in response times in the various conditions.

Additive Factors

The two experimental factors have additive effects on response time. For the digit responses (see TABLE 4),

$$\Delta T_d(\Delta I, \Delta N) = \Delta T_d(\Delta I, 0) + \Delta T_d(0, \Delta N),$$

$$113 \approx 21 + 87,$$

with an error of 5. A similar equation holds for the tone responses

$$182 = 37 + 145,$$

with an error of 0.

The additivity is evidence that the interstimulus interval and the decision about the tone are sequential processes. If all the processes in the task were sequential, though, then when I is prolonged, each response time should increase by however much I increases. But when I is prolonged by 100, T_d is only increased by 21 and T_n by 37. What happened to the rest of the time by which I was prolonged?

A solution is to assume that some process, or sequence of processes, is executed concurrently with I and takes longer than I , as illustrated in FIGURE 4, where U indicates this sequence. In the figure, o_d and o_n are the onsets of the digit and tone, respectively. If the lower path takes longer than the upper path by about 79, then when I is prolonged by 100, the first 79 are expended in simply making the upper and lower paths equal in duration, and the remaining 21 yield the observed increase in T_d .

The lower path U of processes concurrent with I is drawn as preceding N in FIGURE

TABLE 4. Changes in Response Times in Becker's Digit and Tone Identification Task¹⁸

Factor Levels		Prolongations	ΔT_d	ΔT_n
ΔI	Tone Responses			
190	2	$(\Delta I, \Delta N)$	113	182
90	2	$(0, \Delta N)$	87	145
190	1	$(\Delta I, 0)$	21	37
90	1	Baseline	0	0

NOTE: The baseline digit response time was 346, and the baseline tone response time was 475.

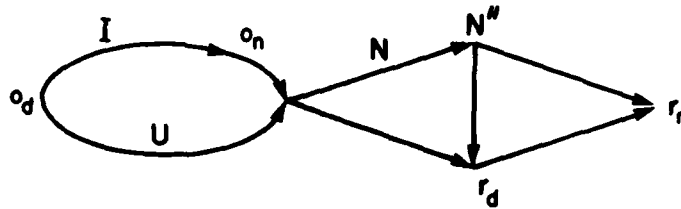


FIGURE 4. A network for Becker's digit and tone identification task. The digit is presented at o_d and the tone at o_n . The interstimulus interval is I and the decision about the tone is N . Responses to the digit and tone are made at r_d and r_n , respectively.

4 for the following reason: Since the effects of prolonging I and N are additive for both reaction times, $k_d(IN)$ and $k_n(IN)$ are both zero, that is, the slack for I with respect to each response. The network in FIGURE 4 is the simplest way of achieving these equalities.

Since U begins before the tone has been presented, U is probably involved in processing the digit. A role for U is suggested by Welford's^{19,20} single-channel theory, which says that when a subject is presented with two stimuli, he can only make a decision about one of them at a time, although peripheral processing of the two stimuli might overlap. A clear explication of this theory is provided in an excellent review by Kerr.²¹ It may be, then, that U involves the decision about the digit, and must be completed before the decision about the tone can start.

The Order of I and N

Since I and N are on a path together, Equation 3 can be used to establish their order. Common sense suggests that I precedes N , since the decision about the tone can only occur after I has ended and the tone is presented. If, contrary to this idea, N precedes I , we would expect Equation 3 to hold in the form

$$\Delta T_d(\Delta I, \Delta N) - \Delta T_d(\Delta I, 0) = \Delta T_n(\Delta I, \Delta N) - \Delta T_n(\Delta I, 0),$$

but the error would be 53 (see TABLE 4), so N does not precede I .

On the other hand, if I precedes N we would expect Equation 3 to hold in the form

$$\Delta T_d(\Delta I, \Delta N) - \Delta T_d(0, \Delta N) = \Delta T_n(\Delta I, \Delta N) - \Delta T_n(0, \Delta N),$$

and from TABLE 4,

$$113 - 87 = 182 - 145.$$

Here the error is only 11, so the evidence supports the common sense order I then N .

Many more details about the network can be determined; in particular, several relationships between the durations of paths in the network can be derived, the details of which are in the appendix.

DISCUSSION

Factorial experiments on reaction time are commonplace due to the influence of Sternberg's¹ additive factor method. But it is rare for an investigator to use more than

two levels of a factor or to use three or more factors. Each level of a factor which is included doubles the size of the experiment, so naturally such experiments are unattractive unless one is looking for something specific. Such experiments are important for testing network models because of the very specific predictions made in Equations 1-4.

The first three experiments discussed were similar in that each manipulated a factor affecting encoding and a factor affecting the information in the stimulus. The results from all three support the idea that these factors prolong different mental processes, provided that we allow for the possibility that more than one process can be executed at a time. In each of the experiments at least one of the network equations was tested and found to hold. Furthermore, two similar experiments, the one by Blackman and the one by Sternberg, led to the same network representation.

More experiments are needed to test the network models, and to determine the principles underlying their forms. One principle is the single-channel hypothesis that a subject can make only one decision at a time.^{19,20,23} There may be other limitations at work in the peripheral processing. For instance, Fisher²² has provided evidence that in search tasks there are a limited number of processors available—perhaps four—and hence there may be a limit on the number of encoding processes that can be executed concurrently. Discovering the arrangements of processes in cognitive tasks is a challenging problem, and discovering the principles underlying these arrangements will be even more so.

REFERENCES

1. STERNBERG, S. 1969. The discovery of processing stages: Extensions of Donders' method. *In* Attention and Performance: II. W. G. Koster, Ed. North-Holland, Amsterdam.
2. COFFMAN, E. G., Ed. 1976. Introduction to deterministic scheduling theory. *In* Computer and Job-Shop Scheduling Theory. John Wiley, New York, NY.
3. CONWAY, R. W., W. L. MAXWELL & L. W. MILLER. 1967. Theory of Scheduling. Addison-Wesley, Reading, MA.
4. KELLEY, J. E. & M. R. WALKER. 1959. Critical path planning and scheduling. *In* Proceedings of the Eastern Joint Computer Conference, Boston, MA: 160-173.
5. MALCOLM, D. G., J. H. ROSEBOOM, C. E. CLARK & W. FAZAR. 1959. Applications of a technique for research and development program evaluation. *Operations Res.* 7: 646-669.
6. TOWNSEND, J. T. 1972. Some results concerning the identifiability of parallel and serial processes. *Br. J. Math. Statist. Psychol.* 25: 168-199.
7. MCCLELLAND, J. L. 1979. On the time relations of mental processes: An examination of processes in cascade. *Psychol. Rev.* 86: 287-330.
8. SCHWEICKERT, R. 1982. The bias of an estimate of coupled slack in stochastic PERT networks. *J. Math. Psychol.* 26: 1-12.
9. CHRISTIE, L. S. & R. D. LUCE. 1956. Decision structure and time relations in simple choice behavior. *Bull. Math. Biophys.* 18: 89-112.
10. FISHER, D. L. & W. M. GOLDSTEIN. Stochastic PERT networks as models of cognition: Derivation of the mean, variance and distribution of reaction time using order-of-processing (OP) diagrams. *J. Math. Psychol.* In press.
11. MCGILL, W. J. & J. GIBBON. 1965. The general-gamma distribution and reaction times. *J. Math. Psychol.* 2: 1-18.
12. SCHWEICKERT, R. 1978. A critical path generalization of the additive factor method: Analysis of a Stroop task. *J. Math. Psychol.* 18: 105-139.
13. DONDERS, F. C. 1868. [Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool. Tweede Reeks.] II. 92-120. *In* Attention and Performance: II (1969). Translated by W. G. Koster (Ed.). North-Holland, Amsterdam.

14. HICK, W. E. & A. T. WELFORD. 1956. Comments on "Central inhibition: Some refractory observations" by A. Elithorn and C. Lawrence. *Q. J. Exp. Psychol.* 8: 39-41.
15. BLACKMAN, A. R. 1975. Test of the additive factor method of choice reaction time analysis. *Percept. Motor Skills* 41: 607-613.
16. MILLER, J. & R. ANBAR. 1981. Expectancy and frequency effects on perceptual and motor systems in choice reaction time. *Memory Cognition* 9: 631-641.
17. SCHWEICKERT, R. 1980. Critical path scheduling of mental processes in a dual task. *Science* 209: 704-706.
18. BECKER, C. A. 1976. Allocation of attention during visual word recognition. *J. Exp. Psychol. Human Percept. Perform.* 2: 556-566.
19. WELFORD, A. T. 1952. The "psychological refractory period" and the timing of high-speed performance—a review and a theory. *Br. J. Psychol.* 43: 2-19.
20. WELFORD, A. T. 1967. Single-channel operation in the brain. *Acta Psychol.* 27: 5-22.
21. KERR, B. 1973. Processing demands during mental operations. *Memory Cognition* 1: 401-412.
22. FISHER, D. L. 1982. Limited channel models of automatic detection: Capacity and scanning in visual search. *Psychol. Rev.* 89: 662-692.
23. SCHWEICKERT, R. 1983. Latent network theory: Scheduling of processes in sentence verification and the Stroop effect. *J. Exp. Psychol. Learn. Memory Cognition* 9: 353-383.

APPENDIX

Details about the Experiments of Sternberg and Blackman

Digit-Naming Network

Let x' and x'' denote the starting point and terminating point, respectively, of a process x . There are three properties required by any critical path network representing the data in TABLE 1. (a) None of the processes E , D , or Q is on the critical path. (b) The longest path from E'' to the response at r does not contain D'' or Q' . (c) The longest path from the stimulus presentation at o to Q' does not contain E'' or D' .

To see that (a) is true, consider process E . Since $k(ED) = s(Er) - s(ED) = 24$ is positive, the total slack for E is positive, and E cannot be critical. Similar arguments show that D and Q are not critical.

Part (b) follows from the following equation.¹² Suppose process x precedes process y . Then

$$k(xy) = \delta(or) - \delta(oy') - \delta(x''r) + \delta(x''y'), \quad (5)$$

where $\delta(uv)$ is the duration of the longest path between points u and v .

If Q' were on the longest path from E'' to r , then

$$\delta(E''r) = \delta(E''Q') + \delta(Q'r).$$

Then, by Equation 5, it is easy to show that $k(EQ) = k(DQ) > 0$. However,

$$k(EQ) - k(DQ) = -3 - 83 = 0,$$

contradicting the idea that Q' is on the longest path from E'' to r . Similar arguments complete the demonstration of (b) and (c).

Path Durations

Since the reaction time is $\delta(or) = 328$, Equation 5 provides information about the durations of paths. For example, since $k(ED) = 24$,

$$328 - k(ED) = 304 = \delta(oD') + \delta(E''r) - \delta(E''D').$$

The other two coupled slacks lead to similar equations.

More Information about the Network in Becker's Tone and Digit Experiment

Two new equations will be introduced here. First, if process x is prolonged by Δx , then if Δx is smaller than $s(xr)$, no increase in the reaction time occurs. Otherwise, the increase in reaction time is $\Delta x - s(x, r)$. That is,

$$\Delta T(\Delta x) = \max\{0, \Delta x - s(x, r)\}. \quad (6)$$

Second, the total slack for x can be written in terms of path durations as follows:

$$s(xr) = \delta(or) - \delta(ox') - \delta(x) - \delta(x''r). \quad (7)$$

Estimates of Slack

In Becker's experiment the size of the interstimulus interval, I , is known because it was directly controlled by the experimenter. With this information, Equation 6 can be used to estimate the size of the total slack for I . Let r_d and r_n be the points at which the responses to the digit and tone, respectively, are made. By Equation 6,

$$\Delta T_d(\Delta I, 0) = \Delta I - S(I, r_d).$$

Because of the experimental procedure, $\Delta I = 100$. Therefore, $s(I, r_d) = 79$. Similarly, $s(I, r_n) = 63$. Since these two estimates are so close, we will assume the slacks are equal, and use the average, 71, as the estimate,

$$s(I, r_d) = s(I, r_n) \approx 71.$$

Location of the Two Responses

Process N precedes both r_d and r_n because prolonging N prolongs both reaction times. Furthermore, r_d is not on the longest path from N'' , the terminus of N , to r_n . To see this, note that

$$\begin{aligned} \Delta T_d(0, \Delta N) - \Delta T_n(0, \Delta N) &= \Delta N - s(N, r_d) - \Delta N + s(N, r_n) \\ &= \delta(o_d, r_n) - \delta(N'', r_n) - \delta(o_d, r_d) + \delta(N'', r_d), \end{aligned}$$

by Equations 6 and 7. If r_d were on the longest path from N'' to r_n , then

$$\delta(N'', r_n) = \delta(N'', r_d) + \delta(r_d, r_n),$$

so

$$\Delta T_d(0, \Delta N) - \Delta T_n(0, \Delta N) = \delta(o_d, r_n) - \delta(o_d, r_d) - \delta(r_d, r_n) \geq 0.$$

But $\Delta T_d(0, \Delta N) - \Delta T_n(0, \Delta N) = 87 - 145 = -58 < 0$, contradicting the idea that r_d is on the longest path from N'' to r_n . Therefore, there is a path from the end of N to r_n , not containing r_d (see FIGURE 4).

Since the subjects were instructed to respond to the digit first, then the tone, I indicate that r_d precedes r_n ; this is not necessary to explain the data, however.

I will now show that (a) the longest path from o_d to r_d does not contain N , and (b) the longest path from o_n to r_d does not contain N . I have shown in the equations above that $s(N, r_d) - s(N, r_n) = 58$, so $s(N, r_d)$ is greater than zero, and there must exist a path from o_d to r_d not containing N . Hence, (a) is true. I have also shown above that $s(I, N) = s(I, r_d)$ and by Equation 6 it can be shown that

$$\delta(o_n, r_d) - \delta(o_n, N') = \delta(o_d, r_d) - \delta(o_d, N').$$

Substituting the left hand side of the above expression into the following expression (Equation 7),

$$s(N, r_d) = \delta(o_d, r_d) - \delta(o_d, N') - \delta(N) - \delta(N'', r_d) > 0,$$

we obtain

$$\delta(o_n, r_d) - \delta(o_n, N') - \delta(N) - \delta(N'', r_d) > 0,$$

so (b) is true. Perhaps the simplest way to incorporate propositions (a) and (b) into the model is with the network of FIGURE 4.

In order to account for all the time elapsing between the onset of the first stimulus and the onset of the last response, one would like to know the duration of each process in the network. It is not possible with these data to completely determine the durations, but we can come close. The duration of the longest path between every two points in the above network can be expressed in terms of four parameters, $\delta(o_n, N')$, $\delta(N)$, $\delta(N'', r_d)$ and $\delta(r_d, r_n)$, and an upper bound can be found for each parameter. The details are left to the reader.

Timing Perturbations with Complex Auditory Stimuli^a

DONALD G. JAMIESON, ELZBIETA SLAWINSKA,
MARGARET F. CHEESMAN, AND
BLAS ESPINOZA-VARAS

*Department of Psychology
The University of Calgary
Calgary, Alberta, T2N 1N4 Canada*

In a duration-discrimination task, the order in which stimuli are presented can have a substantial effect on accuracy. For example, subjects are more often able to select the longer of a pair of durations, say 340 msec followed by 300 msec, when the stimuli are presented in that order (that is, longer followed by shorter) than in the reverse order (300 followed by 340).^{1,2} Such a presentation order effect is called a positive time-order error (TOE). Time-order errors are not unique to duration comparisons, of course, since they have been reported with a number of stimulus modalities, including brightness and loudness, but they are particularly sizable with durations. For example, for brief-duration stimuli, such as those described above, the time-order error may produce a difference in the proportion of correct responses between the two presentation orders of 0.3, or more.³

Positive TOEs are thus well documented for the case of judgments of the overall duration of pairs of stimuli.^{2,4} In other situations, when the overall set of stimuli covers a substantial range, it is well established that another form of presentation order effect—the assimilation effect—is likely to occur when subjects judge overall duration.^{1,2,5} In a number of interesting perceptual situations, however, the listener must judge the duration of components or segments of complex stimuli. In one class of such complex stimuli, speech sounds, the duration of an acoustic event may provide an important cue to phonetic class. In fact, one such distinction—voice-onset time (VOT)—is arguably the most commonly studied variable in speech perception research.⁶⁻⁸ VOT discrimination between /ba/ and /pa/ requires, among other things, attention to the duration of events occurring within the first 80 to 100 msec of a syllable that is 300 to 400 msec in length. It is therefore relevant to ask (a) whether TOE and/or assimilation effects occur for such “duration judgments,” and (b) how they may compare with the effects known to occur with judgments of overall stimulus duration.

The three experiments reported in this study sought to answer these questions. The first experiment examined judgments with synthetic /ba/-/pa/ syllables differing in VOT. The second experiment examined judgments with pure-tone stimuli consisting of a frequency glide plus a steady-state tone. Here the variable of interest was the duration of the initial transition. The third experiment examined judgments of the identity of two tones containing differences in the duration and/or in the frequency of one tone of the pair. These experiments also allow us to examine the influence of the temporal and spectral complexity of the stimuli on the TOE effects.

^aThis work was supported by grants from the Natural Sciences and Engineering Research Council and the Alberta Heritage Foundation for Medical Research (to D.G.J.)

Method

Our procedure involved a "same-different" task, in which a standard stimulus was selected from the continuum, then followed by a comparison stimulus, also from the continuum. On 50% of the trials, the standard and comparison stimuli were physically identical; on the remaining trials, the two stimuli differed by one step of 10 msec. The 10-msec difference could consist of either an increment or a decrement in VOT. Thus, on one-half of the different trials, the comparison stimulus was longer than the standard and on the other half, it was shorter than the standard. The task required the subject to indicate whether the two stimuli presented on any trial were "the same," or whether they were "different," by depressing one of two labeled response buttons. Stimulus presentation, response collection, and all aspects of timing and sequencing were controlled by a PDP11/34 computer.

During testing, the subject was seated comfortably in an IAC sound-attenuating chamber. Stimuli were presented to the right ear at 72 dB SPL, measured by a Bruel and Kjaer type 2118 sound-level meter, coupled to a type 4152 artificial ear. All three subjects had normal hearing, and previous experience listening to synthetic speech.

Subjects served for a total of four 140-trial experimental sessions. Each trial consisted of the following sequence of events: an intertrial interval of 1000 msec preceding the first stimulus; the presentation of the first stimulus; a silent, 400-msec interstimulus interval; the presentation of the second (comparison) stimulus; a response interval of 4 seconds.

Results and Discussion

The traditional measure of the time-order effects which occur in duration comparison situations is obtained by comparing discrimination scores in the two presentation orders: "shorter, longer" (that is, when the shorter duration precedes the longer), denoted (*S, L*), and "longer, shorter," denoted (*L, S*). The statistic, $TOE = P(C|L, S) - P(C|S, L)$, is zero when there is no time-order error. It is greater than zero when the time-order error is "positive," and less than zero when the time-order error is "negative."

Since our interest is in the possibility of presentation order effects with these stimuli, we will confine our analysis to the time-order error index, $TOE = P(C|L, S) - P(C|S, L)$, where *L* indicates the longer VOT interval (that is, more pa-like stimulus), *S* indicates the shorter VOT interval (that is, more ba-like stimulus), and *C* indicates that the correct discrimination (different) was made. Thus, $P(C|L, S)$ indexes the proportion of correct responses observed when the longer duration of a stimulus pair was the first presented.

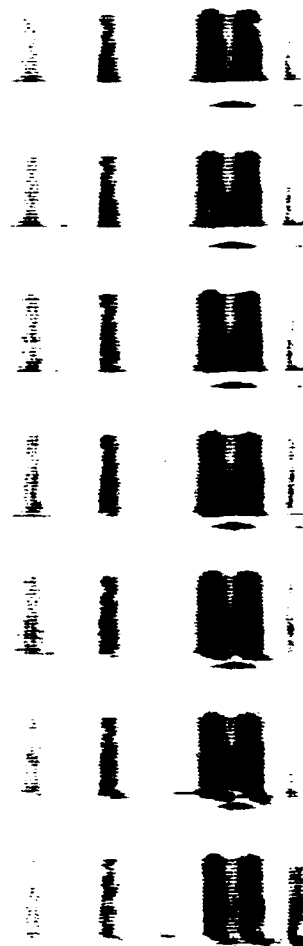
FIGURE 2 presents the value of this index as a function of the duration of the longer VOT of the pair for a typical subject. A clear trend is apparent in the data: the TOE index is uniformly positive for durations at the /ba/ end of the continuum. That is, a comparison VOT, 10 msec shorter than the standard is discriminated better than a comparison VOT 10 msec longer than the standard. This result is consistent with the notion that positive time-order errors should occur when brief durations are compared. However, the TOE index decreases regularly with increasing VOT duration, becoming negative with the longer VOT duration at the /pa/ end of the continuum. This latter result is more consistent with the occurrence of assimilation effects² than with positive TOEs.

It is thus clear that the two fundamental presentation order effects observed when

EXPERIMENT 1: PRESENTATION ORDER EFFECTS WITH SPEECH STIMULI VARYING IN VOICE-ONSET TIME

In articulatory terms, voice-onset time is the interval of time separating the release of the stop from the onset of vocal cord vibration. Acoustically, the release is indicated by a brief noise burst, and vocal cord vibration is indicated by periodicity. Synthetic consonant vowel syllables varying in VOT differ in terms of the delay between onset of the burst and the onset of the periodicity. For "voiceless" consonant-vowel syllables

FIGURE 1. The seven synthesized speech stimuli used in Experiment 1. Stimuli towards the *left* have brief VOT values, and are heard as /ba/. Stimuli towards the *right* have longer VOT values, and are heard as /pa/.



(CVs), such as /pa/ and /ta/, voice-onset time values may reach 80-100 msec; for "voiced" CVs, such as /ba/ and /da/, voice-onset time values are less than 20 msec.

Stimuli

A seven-stimulus voice-onset time continuum, ranging from /ba/ to /pa/, was synthesized using a digital speech synthesizer⁹ implemented on a PDP11/34. The resulting stimuli are presented in FIGURE 1.

overall stimulus duration is compared are also observed with comparisons of only a portion of a complex acoustical stimulus. First, time-order errors occur even when the perceptual judgment refers to the duration of the initial events of a longer acoustic complex. Second, assimilation effects clearly occur in such a situation. As well, these results indicate that speech stimuli are susceptible to two of the "nuisance" variables of psychophysics, extending a result established with a selective adaptation speech paradigm.⁷

EXPERIMENT 2: DO PRESENTATION ORDER EFFECTS OCCUR WITH ANALOGUES TO SPEECH FORMANT TRANSITIONS?

The results of the previous experiment were obtained with stimuli that were both acoustically complex and capable of inducing a phonetic percept. Experiment 2

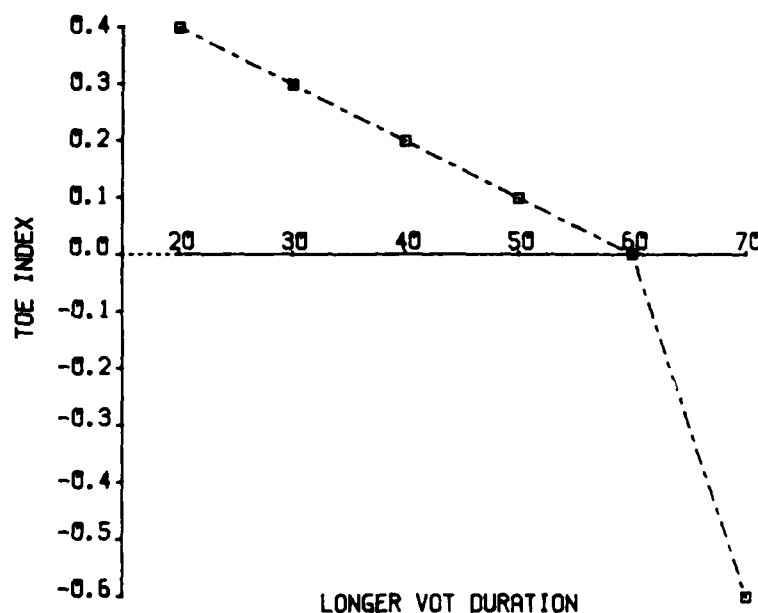


FIGURE 2. The time-order error index, $TOE = [P(C|L, S) - P(C|S, L)]$, for a typical subject, as a function of the longer VOT duration. In each case stimuli are adjacent speech stimuli on a synthesized /ba/ /pa/ VOT continuum. Thus, the longer duration always exceeds the shorter by 10 msec.

eliminated the phonetic component, while retaining some of the complex, time-varying properties of speech sounds. To accomplish this, we used stimuli consisting of an initial frequency glide followed by a steady-state portion, which resembled the variations in first-formant center frequency of a /ba/ sound, but were not heard as speech. These stimuli also have potential importance because such frequency transitions cue important phonetic distinctions such as place of articulation in stop consonants.⁸ As well, researchers seeking to study the auditory precursors to speech perception have seized on the importance of sensitivity to frequency transitions, offering evidence that sensitivity to frequency change exists in distinct "frequency channels."⁹

Stimuli

The stimuli consisted of an initial frequency glide ranging from 400 to 700 Hz, followed by a steady-state tone of 550 Hz. The stimuli differed in terms of the rate of the frequency change of the initial glide or, equivalently, in transition duration. The 300-Hz frequency increment occurred over 20 msec for the fastest change, and over 80 msec for the most gradual change. The seven stimuli covered this range in 10-msec steps (that is, in 20, 30, 40, 50, 60, 70, and 80 msec). In each case, the steady-state portion of the stimulus was adjusted to fix overall stimulus duration at 200 msec. Listeners described these stimuli as "chirps" or "whistles"; they were not heard as speech-like by any listener.

Method

The testing procedure used the same-different paradigm, described for Experiment 1. All aspects of the testing situation were as described for that experiment.

Results and Discussion

The TOE index, $TOE = P(C|L, S) - P(C|S, L)$, where L indicates the stimulus with the longer glide duration and S indicates the shorter glide duration, was computed for each stimulus pair and plotted as a function of the duration of the longer glide. FIGURE 3 shows that this index is positive for the three briefest glide durations, but it becomes negative for the two longer glide durations. Improved performance is seen

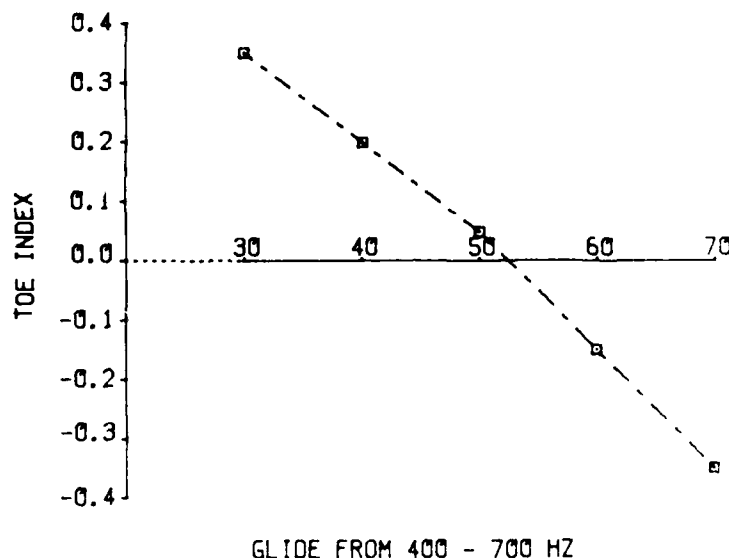


FIGURE 3. The time-order error index, $TOE = [P(C|L, S) - P(C|S, L)]$, for a typical subject, as a function of the longer glide duration. In each case stimuli are pairs of complex sounds, each consisting of a frequency glide from 400 to 700 Hz, followed by a steady-state pure tone at 550 Hz. In each pair, the longer glide duration exceeds the shorter by 10 msec. Overall stimulus duration was fixed at 200 msec in each case; stimuli towards the *left* of the figure have more rapid glides, while stimuli towards the *right* have slower glides.

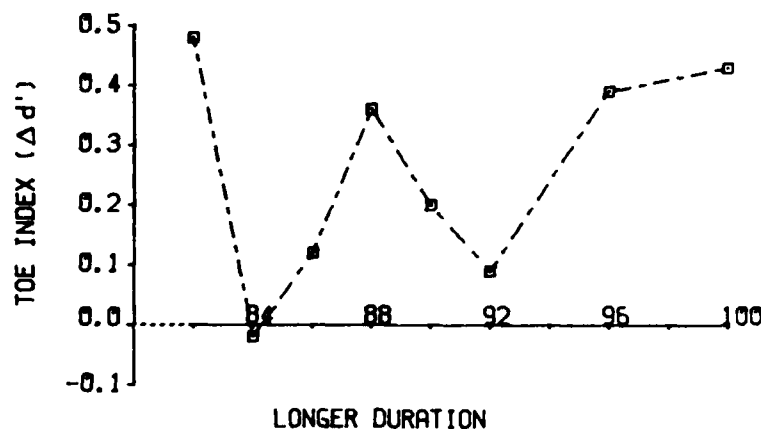


FIGURE 4. The time-order error index, $TOE = [(d'|L, S) - (d'|S, L)]$, where L is the longer duration of the pair, S is the shorter duration, and d' is the familiar measure of discriminability. Note that the shorter duration is fixed at 80 msec in each case.

when the second-presented stimulus was the more extreme of the pair and not routinely when the longer duration was the first presented. There is thus little or no evidence for a dominant, or even important positive time-order error. Rather, assimilation appears to be the major effect in this experiment.

The failure to find a substantial positive TOE in Experiment 2 limits the generality of the conclusions of Experiment 1. Positive TOEs may thus occur under two types of conditions: in comparisons of the overall duration of simple stimuli, and in comparisons of the duration of cues occupying a portion of sufficiently complex stimuli. In this view, the stimuli of Experiment 2 would be insufficiently complex; a further experiment using still simpler "complex" stimuli would not show positive TOEs.

The consistent occurrence of a substantial assimilation effect in Experiments 1 and 2 is also of interest, of course. Analyses of performance with speech stimuli and with speech-analogue stimuli rarely examine the possibility of such effects. As seen here, however, these effects can be substantial in magnitude, resulting in differences in the proportion of correct responses of as much as 40% between presentation orders. Assimilation effects accompany the use of a substantial stimulus range, however. To eliminate such effects in Experiment 3, we restricted the overall stimulus range substantially, relative to the preceding experiments.

EXPERIMENT 3: PRESENTATION ORDER EFFECTS WITH SPEECH ANALOGUES TO SPECTRAL/TEMPORAL FUSION

Stimuli

This experiment used pairs of pure-tone stimuli which (a) were identical 80-msec 1500-Hz sounds or (b) differed in the frequency of one tone and/or in the duration of the other tone. Stimuli were generated by analogue equipment and presented at 75 dB over matched and calibrated TDH-49 headphones.

Method

The experimental procedure was essentially as described for previous experiments, except that the interstimulus interval was 60 msec and the two conditions in which (a)

the first tone of the pair could be incremented in duration and (b) the second tone of the pair could be incremented in duration were run sequentially as different experiments.

Results and Discussion

Performance for each stimulus in the two experimental conditions was first summarized by the d' statistic. The TOE index, $TOE = (d' | L, S) - (d' | S, L)$, where L indicates the longer duration of the pair and S indicates the shorter duration of the pair (which was fixed at 80 msec), and d' is the familiar discrimination measure of signal detection theory, is presented in FIGURE 4. It is clear from this figure that discrimination scores tend to be higher when the first-presented duration is the longer, consistent with the expectation based on explicit duration-comparison experiments.

GENERAL DISCUSSION

The present experiments show that positive time-order errors and/or assimilation effects occur when brief-duration stimuli are compared, even when duration is but a cue in part of a complex auditory stimulus. As such, these results have potential implications for experimenters whose major interest is far removed from the study of timing and duration perception. Of course, since these results suggest that such phenomena are by no means unique to the explicit duration-comparison situation, they may be considered encouraging to traditional duration-perception research.

REFERENCES

1. ALLAN, L. 1977. The time-order error in judgments of duration. *Can. J. Psychol.* **31**: 24-31.
2. JAMIESON, D. 1977. Two presentation order effects. *Can. J. Psychol.* **31**: 184-194.
3. JAMIESON, D. & W. PETRUSIC. 1976. On a bias induced by the provision of feedback in psychophysical situations. *Acta Psychol.* **40**: 199-206.
4. JAMIESON, D. & W. PETRUSIC. 1975. Presentation order effects in duration discrimination. *Percept. Psychophys.* **17**: 197-202.
5. HELLSON, H. 1964. *Adaptation-level theory*. Harper & Row, New York, NY.
6. ODEN, G. & D. MASSARO. 1978. Integration of featural information in speech perception. *Psychol. Rev.* **85**: 172-191.
7. DIBHI, R., M. LANG & E. PARKER. 1980. A further parallel between selective adaptation and contrast. *J. Exp. Psychol. Human Percept. Perf.* **6**: 24-44.
8. WOOD, C. 1976. Discriminability, response bias and phonemic categories in the discrimination of voice onset time. *J. Acoust. Soc. Am.* **60**: 1381-1388.
9. KLATT, D. 1980. Software for a cascade/parallel formant synthesizer. *J. Acoust. Soc. Am.* **67**: 971-995.
10. RIPP, B., A. HEALEY & R. CROWDER. 1979. Categories and context in the perception of steady state vowels. *J. Exp. Psychol. Human Percept. Perf.* **5**: 129-145.
11. REGAN, D. & B. TANSLEY. 1979. Selective adaptation to frequency-modulated tones. *J. Acoust. Soc. Am.* **65**: 1249-1257.

Temporal Order and Duration: Their Discrimination and Retention by Pigeons

EDWARD A. WASSERMAN, ROBERT E. DELONG, AND
MARK B. LAREW

*Department of Psychology
The University of Iowa
Iowa City, Iowa 52242*

The external environment contains a vast variety of energies which organisms may detect with specialized receptor organs. Beyond these highly specific sensory systems, however, organisms may perceive properties of single and multiple stimuli that are in some measure independent of the particular form of energy that is acting upon the sensory receptor.^{1,2}

For instance, consider a patch of colored light projected on a nickel-sized disk. A pigeon may readily discriminate the hue, brightness, and shape of the patch. It may also discriminate how long the light has remained on. To the extent that duration is a discriminable property of the visual stimulus for the bird, we have reason to believe that the pigeon would similarly discriminate the duration of nonvisual stimuli, such as sounds, odors, and tastes. In short, duration is a property of every sensory stimulus. What we learn about duration discrimination in one modality ought to hold in another modality (for a possible exception see Spetch and Wilkie³).

Now, consider the successive presentation of a pair of stimuli. The two stimuli may either be the same or different. And, if they are different, they may occur in one of two temporal orders. Of course, to discriminate whether two stimuli are the same or different as well as to discriminate the temporal order of two nonidentical stimuli, sensory receptors must be stimulated. But here again, we expect that same-different and temporal-order discriminations ought to occur in all remaining sensory modalities, given successful demonstration in one. In fact, same-different and temporal order discriminations are so general that they ought to hold *across* as well as *within* all sensory systems.

Our interest in duration and temporal order discrimination has prompted a series of experiments using operant conditioning techniques and pigeon subjects. This work has persuaded us that pigeons very ably discriminate these features of visual stimuli. In addition, by inserting delays between the discriminative stimuli and the performance test, we have traced the retention of these discriminations over the course of several seconds. Duration and temporal order are thus both discriminable and rememberable stimulus properties for our avian subjects.

TEMPORAL ORDER

Our investigations into temporal-order discrimination⁴⁻⁶ were initially inspired by the aim of definitively demonstrating control over behavior by this aspect of two nonidentical stimuli. As we pondered the problem, several possible paradigms fell short of our goal. Suppose, for example, that as stimulus pairs we were to use colored key lights ordered orange-green and green-orange. Were we to reinforce the pigeon's key

pecks to a vertical-line test stimulus after orange-green pairs but not to do so after green-orange pairs, we would expect test-key pecking in the first case to exceed that in the second. But would we have clearly shown that stimulus order was the critical aspect of the stimulus pairs that was the basis for the bird's discrimination? No. The problem here is that selectively attending to either the first or to the second color of the pair would be sufficient for the pigeon to respond discriminately on orange-green and green-orange trials; for instance, to peck the vertical line if the second color was green, but not if it was orange.

One solution to this difficulty would be to intermix orange-orange and green-green pairs with trials involving orange-green and green-orange pairs. As before, pecks to the vertical-line test stimulus would result in food only after orange-green pairs. But now, selective attention to either the first or to the second color of the pair would fail to provide the bird with a basis for accurate discriminative responding.

Even with these modifications, pigeons can learn to respond at high rates on orange-green tests and at low rates on green-orange, orange-orange, and green-green tests (see experiments 1 and 2 in Weisman *et al.*⁶ for such data with slightly different visual stimuli). This performance clearly indicates that a sequence of two colors presented in a specific temporal order can acquire discriminative control over a pigeon's behavior. Yet there is still a means by which the pigeon could respond discriminatively during the test periods and have no direct access to the identities of the prior color stimuli or to their temporal order. The bird might simply peck the key if the first color were orange and continue to peck into the test if the second color were green; if the first color were green or the second color were orange, the bird might become unresponsive or direct its behaviors away from the response key. Stepwise discrimination of this sort may well afford the pigeon a simple mode of solving this complex task.

A basic problem with the second discrimination task is then that, because only the orange-green sequence is associated with reinforcement, the reversely ordered green-orange sequence can potentially be discriminated as involving nonreinforcement as soon as green is presented as the first color. Therefore, there is no assurance that the pigeon attends to both colors in a sequence and then responds differentially to the test stimulus on the basis of the temporal order of the preceding color pair.

Our solution to this problem was to devise the task outlined in TABLE 1. This task is a variant of a delayed, go/no-go conditional discrimination.^{8,11} Here, orange-green and green-orange stimulus orders are associated with either reinforcement or nonreinforcement, depending upon which of two line orientations, vertical or horizontal, is

TABLE 1. Procedure Used in the Study of Temporal Order Discrimination

Trial Type	Color 1 (2.0 sec)	Trial Sequence				Trial Outcome (2.5 sec)	Intertrial Interval (15.0 sec)
		Interitem Interval (0.5 sec)	Color 2 (2.0 sec)	Retention Interval (0.5 sec)	Test Stimulus (5.0 sec)		
AA	Orange		Orange		Vertical	Blackout	
BB	Orange		Orange		Horizontal	Blackout	
AB	Orange		Green		Vertical	Food	
BA	Orange		Green		Horizontal	Blackout	
BA	Green		Orange		Vertical	Blackout	
AB	Green		Orange		Horizontal	Food	
BB	Green		Green		Vertical	Blackout	
AA	Green		Green		Horizontal	Blackout	

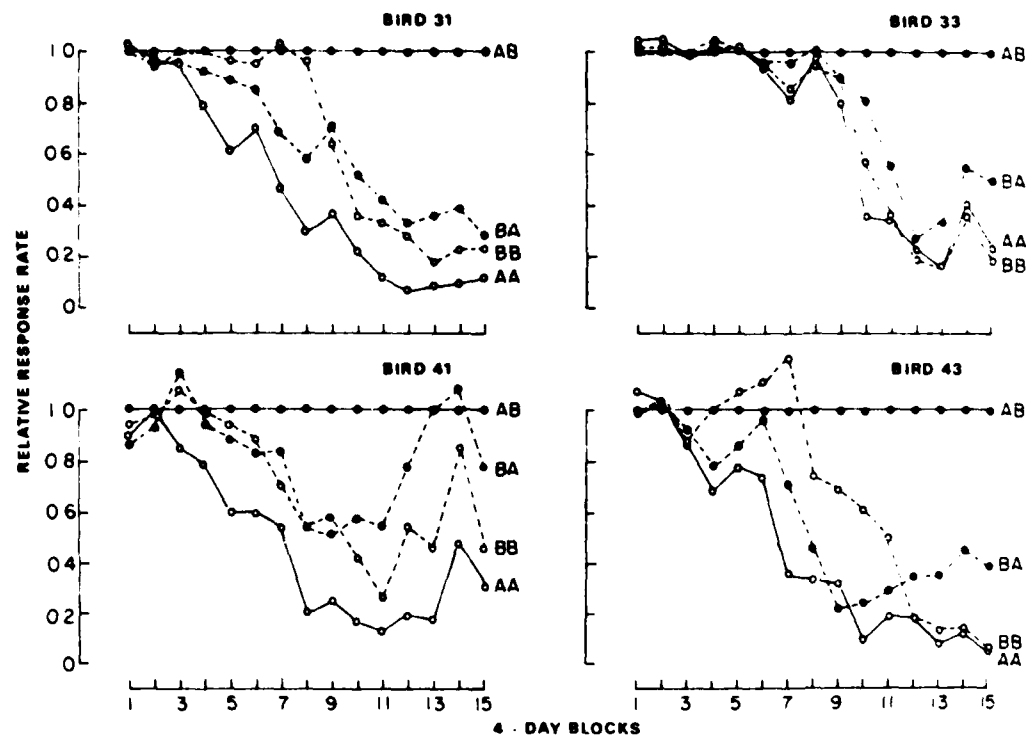


FIGURE 1. Relative rates of responding on *AB*, *BA*, *AA*, and *BB* test trials as a function of 4-day blocks of training. Relative response rates were calculated by dividing a given test rate of response by that obtained on *AB* test trials. (Modified from Weisman *et al.*⁶)

presented as the test stimulus. As can be seen in TABLE 1, when an orange-green pair is followed by a vertical line or a green-orange pair is followed by a horizontal line (designated as *AB* trials in the first column), pecking to the test stimulus is reinforced by food. Alternatively, when an orange-green pair is followed by a horizontal line or a green-orange pair is followed by a vertical line (designated as *BA* trials), pecking to the test stimulus is followed by blackout. Thus, the order of color pairs associated with reinforcement is conditional upon the test stimulus. This conditionality sought to require the bird to remember the order of the colors until the test stimulus was presented, because a particular order of the two nonidentical colors could not be discriminated as reinforced or nonreinforced until that time. Note, however, that the conditionality held only for trials involving nonidentical colors; trials involving pairs of identical colors never entailed reinforcement. Here, *AA* and *BB* refer to trials with identical colors in which the prevailing color corresponds, respectively, to the first or to the second item of *AB* trials for a particular test stimulus.

Discrimination

To examine the acquisition of discriminative responding under the procedures of TABLE 1, four pigeons were trained for 60 days, with key pecking separately recorded on *AB*, *BA*, *AA*, and *BB* tests. Daily sessions comprised 80 trials, with each of the eight possible sequences occurring randomly and equiprobably. FIGURE 1 (from experiment 3, replication 1 of Weisman *et al.*⁶) plots responding on all trial types relative to

responding on *AB* trials, over successive 4-day blocks of training. Such a depiction, of course, means that scores on *AB* trials always equal 1.0.

In general, discrimination training resulted in a gradual reduction in erroneous responding on nonreinforced test trials (*BA*, *AA*, and *BB*), although the performance of bird 41 deteriorated from its prior high level near the end of training. As to the relative rates of discriminating reinforced from nonreinforced trial types, *AA* trials were most readily discriminated from *AB* trials; *BA* and *BB* trials were discriminated from *AB* trials with greater difficulty. Furthermore, whereas early in training erroneous responding on *BB* tests equalled and sometimes exceeded responding on *BA* tests, later in training erroneous responding on *BA* tests usually exceeded that on *BB* tests.

Clearly, even under the demanding procedure that we devised, pigeons were able to discriminate a recently presented pair of nonoverlapping, nonidentical color stimuli from (a) the same two colors presented in the opposite temporal order and (b) stimulus pairs comprising two identical colors.

The pattern of differential responding over the course of training suggests that final performance followed mastery of two discriminations. First, responding appeared to be based mainly on the second color of the pair and its relationship to the line orientation test stimulus. Although both *A* and *B* colors consistently preceded the test stimuli on reinforced trials, *B*-test associations exerted stronger control over differential responding than did *A*-test associations; thus, response rates on *AB* and *BB* tests initially exceeded those on *BA* and *AA* tests. Second, the birds successfully distinguished *AB* trials from all others. Correlated with this discrimination was a tendency for erroneous responding on *BA* tests to exceed that on *AA* and *BB* tests. This result may be due to the subjects' discriminating trials with identical colors from those with nonidentical colors; only the latter could eventuate in food reinforcement.

Retention

Given clear evidence of stimulus-order discrimination, we were interested in seeing how long such information could be retained. Studies of short-term memory in the pigeon using either choice¹² or go/no-go¹³ procedures typically disclose that discriminative performance is a negative function of the sample-test retention interval. Here, we were concerned with determining whether a similar loss of discriminative control would also be observed in our delayed temporal-order discrimination. In particular, the possibility of the pigeon's discriminating the nonoccurrence of reinforcement after pairs of identical colors made it likely that a rather different retention function would be observed following these color pairs than after pairs of nonidentical colors, in which either reinforcement or nonreinforcement were equiprobable outcomes.

Three of the pigeons (birds 31, 33, and 41) continued in this investigation. Daily sessions comprised 96 trials, 12 randomized blocks of the eight possible trial sequences shown in TABLE 1. The first 16 trials in each session were warm-up trials, and the retention interval here was always 0.5 sec; the duration of the retention interval for the final 80 trials was varied between sessions. Retention intervals of 0.5, 1.0, 2.0 and 4.0 sec were randomly presented within 4-day blocks for 16 days. Bird 41 subsequently received 16 additional daily sessions with retention intervals of 0.5, 4.0, 8.0, and 12.0 sec.

The top portion of FIGURE 2 shows that, at the shortest retention interval of 0.5 sec, all three pigeons discriminated reinforced from nonreinforced trials; response rates on *BA*, *BB*, and *AA* tests were all less than half the rate on *AB* tests. Again, response rates were generally ordered *AB* > *BA* > *BB* > *AA*. This ordering is consistent with the

birds' discriminating both identical from nonidentical color pairs ($AB + BA > BB + AA$) and earlier from later items in reinforced color pairs ($AB + BB > BA + AA$). Increasing the retention interval from 0.5 to 4.0 sec substantially disrupted discriminative performance on *BA* trials; however, on *BB* and *AA* tests relative response rates remained at or below 0.4 of the rate on *AB* tests.

The generally poorer discrimination on *BA* than on *BB* and *AA* trials at the shortest retention interval plus the greater loss of stimulus control on *BA* than on *BB* and *AA* trials as the retention interval was lengthened are results consistent with different discriminative and retentive processes operating on trials with identical and nonidentical color pairs. Further support for this proposal is given in TABLE 2. This table shows the rate of key pecking of each pigeon to both color items during each of the eight

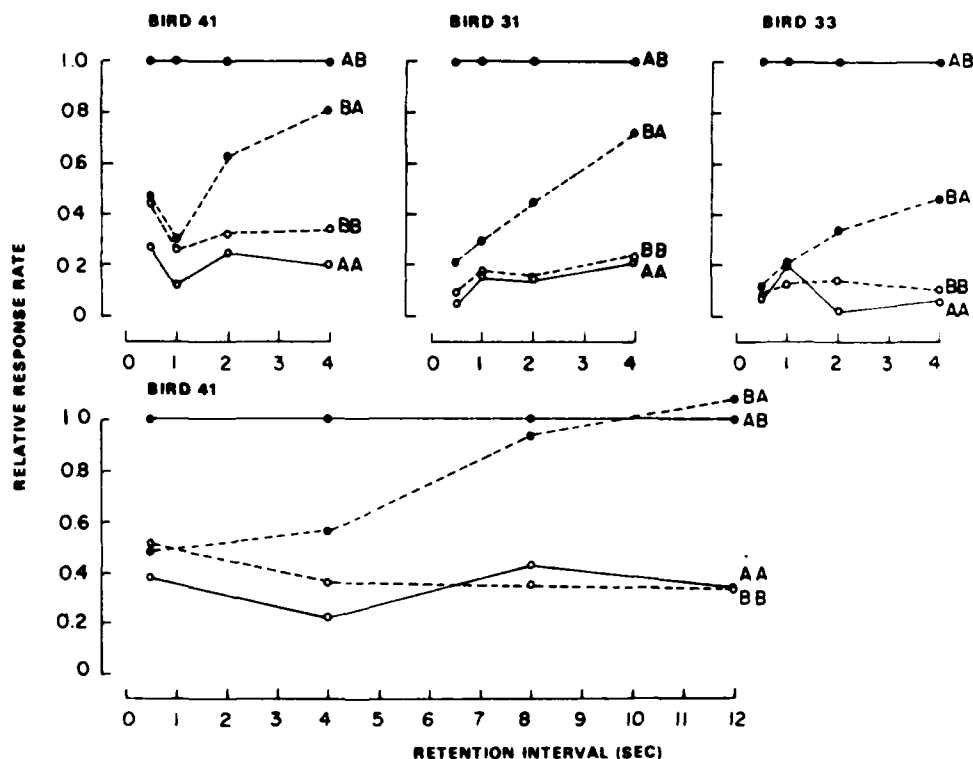


FIGURE 2. Relative rates of responding on *AB*, *BA*, *AA*, and *BB* test trials as a function of the retention interval.

possible color pair-test sequences at the 0.5-sec retention interval. Two trends can be discerned. First, on trials with nonidentical colors, responding to the second item exceeded that to the first. Second, and more pertinent to the present issue, response rates during the second color were much lower when it was the same as the first than when the two colors were different. Thus, the pigeons did indeed discriminate identical from nonidentical colors during presentation of the second item of the pair.

Discrimination of impending nonreinforcement during the second color on *AA* and *BB* trials could be of considerable importance in understanding why little or no loss in discriminative control was observed here compared with that of *BA* trials. Were it easier for pigeons to remember the expectancy of nonreinforcement (on *BB* and *AA*

TABLE 2. Mean Rate of Key Pecking (Pecks per Sec) during Color 1 and Color 2 of Each Color Pair-Line Combination at the 0.5-Sec Retention Interval for Birds 31, 33, and 41

	Color Pair-Line Combinations							
	O-O-V	O-G-V	G-O-V	G-G-V	O-O-H	O-G-H	G-O-H	G-G-H
Bird 31								
Color 1	2.16	2.08	2.09	2.03	2.25	2.20	2.08	2.24
Color 2	0.82	2.67	2.78	0.72	0.62	2.73	2.25	0.37
Bird 33								
Color 1	1.43	1.13	1.33	1.73	1.59	1.79	1.38	1.79
Color 2	0.19	2.08	2.28	0.39	0.30	2.14	2.25	0.48
Bird 41								
Color 1	0.57	0.45	0.60	0.54	0.63	0.23	0.67	0.47
Color 2	1.04	1.65	1.40	1.02	1.18	1.64	1.63	0.93

NOTE: Italicized rates denote responding to color 2 on trials involving nonidentical stimuli.
 ABBREVIATIONS: O = orange; G = green; V = vertical; H = horizontal.

trials) than the order of presentation of earlier, nonidentical colors (on *AB* and *BA* trials), then the present results would have a plausible explanation.¹⁴

Recall that bird 41 received an extra test series with 0.5-, 4.0-, 8.0-, and 12.0-sec retention intervals. The bottom portion of FIGURE 2 again shows: fairly accurate discrimination of stimulus order at the shortest 0.5-sec retention interval; loss of discriminative control on *BA* trials when the retention interval reached 8.0 and 12.0 sec; and no loss of discriminative control on *BB* and *AA* trials, even at the 12.0-sec retention interval.

We now knew that increasing the retention interval between the second color item and the test stimulus impaired the discrimination between prior *AB* and *BA* orders, but had little effect upon the discrimination of prior *AB* from *AA* and *BB* orders. If such different retention results are indeed due to different memory processes operating *after* the birds discriminate identical from nonidentical color pairs, then varying the interitem interval—which elapses *before* this same-different discrimination can take place—should not differentially affect performance on nonreinforced trials involving identical and nonidentical color pairs. Our final experiment in this series examined the issue by systematically varying both the interitem and the retention interval in a within-sessions design.

Two other pigeons with earlier experience in a related temporal-order task¹⁵ were first pretrained on the problem shown in TABLE 1. The experimental phase of training was conducted in 3-day blocks: one session of baseline, one session of interitem interval manipulation, and one session of retention interval manipulation. The baseline session was the first day of each block; the interitem interval and retention interval manipulation sessions were presented on the second and third days of each block with the order of their occurrence reversed in each successive 3-day block. The experimental phase lasted 24 days.

Each session of experimental training comprised 96 trials. Baseline sessions entailed 12 randomized blocks of the eight sequences outlined in TABLE 1, with the interitem and retention intervals set at 0.5 sec. The first 32 (warm-up) trials of interitem interval manipulation sessions and retention interval manipulation sessions were identical to baseline sessions. The final 64 trials comprised two randomized blocks of 32 trials, with each of the eight stimulus sequences combined with each of four values of the interitem interval or the retention interval: 0.5, 1.0, 2.0, and 4.0 sec.

The primary behavioral measure was the rate of key pecking during the test stimuli, recorded separately for each main trial type (*AB*, *BA*, *AA*, and *BB*). For the final 64 trials of interitem interval manipulation sessions and retention interval manipulation sessions, response rates were also calculated separately for each interval value. Mean response rates were obtained over the eight sessions of each of the two interval manipulations, and relative rates were calculated by dividing rates on all trial types by rates on *AB* trials. (Data from the baseline sessions were not scored since each session of experimental training included trials on which the interitem and retention intervals were both set equal to 0.5 sec.)

Relative response rates of both birds on all trial types are shown in FIGURE 3 as a function of the interitem interval (top portion) and the retention interval (bottom portion). Increasing the retention interval again led to a notable loss of discriminative control on *BA* trials and to a much smaller loss on *AA* and *BB* trials. Bird 12 showed no difference in responding after *AB* and *BA* orders at retention intervals of 1.0 sec or more, whereas bird 42 discriminated between prior *AB* and *BA* orders out to a 2.0-sec retention interval.

Unlike the differential effects of retention interval manipulation on discriminative responding, increasing the interitem interval over the same range of values led to a similar loss of discrimination on *BA*, *AA*, and *BB* trials. For neither bird did

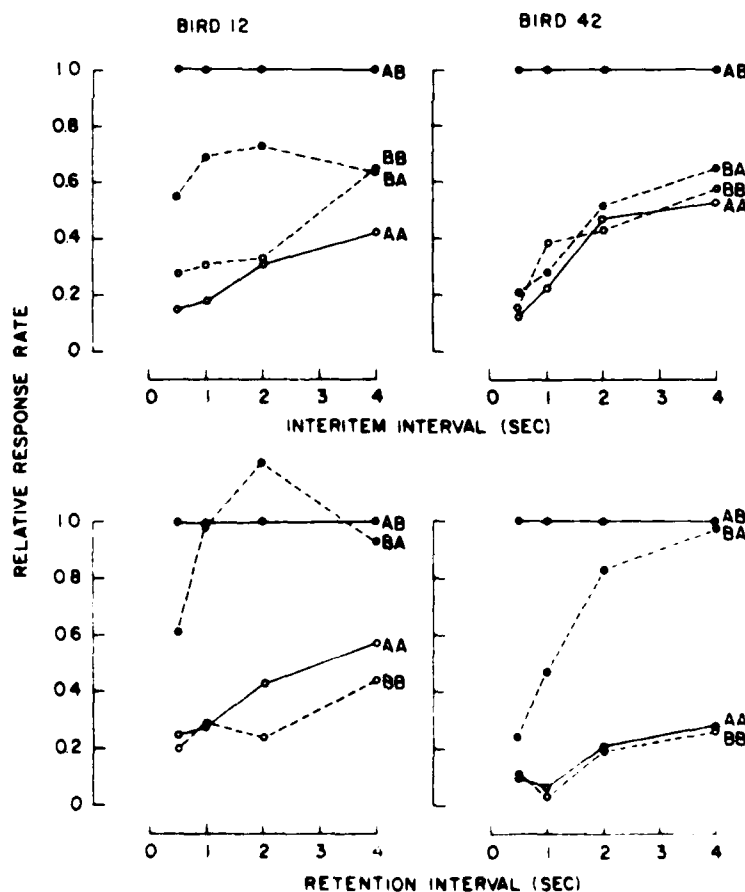


FIGURE 3. Relative rates of responding on *AB*, *BA*, *AA*, and *BB* test trials as a function of the interitem interval (*top*) and the retention interval (*bottom*).

performance on *BA* trials worsen more than on *AA* and *BB* trials. These results thus confirm our suspicion that only *after* presentation of the second color item do different memory processes mediate performance on same-color and different-color trials when the retention interval is lengthened. Because our pigeons could not know in advance whether the second color item would be the same as the first, increasing the interitem interval had equivalent effects on all nonreinforced trial types.

DURATION

That animal behavior can be controlled by the passage of time is obvious when one considers the patterns of responding that are supported by temporal schedules of reinforcement.¹⁵ As noted by Catania,¹⁶ however, such evidence fails to isolate the temporal interval being discriminated from the subject's ongoing stream of behavior. In order to be completely confident in the subject's discriminating two or more stimulus durations, it is necessary to separate the discrimination phase of performance from the report phase.

One rather popular procedure adopting this line of attack proceeds as follows: After pecking a white warning signal, the pigeon is exposed to one of two different durations of a red-lighted key. Following key light offset, two side keys are made available to the pigeon, with left key choices reinforced after one duration and right key choices reinforced after the other duration. If the pigeon were to respond discriminatively after the two stimulus durations, we would have evidence that duration is functioning as a discriminative dimension of the pigeon's environment.

But just how good is this evidence? At least one behavioral strategy can solve this problem without requiring the pigeon to discriminate *each* of the programmed stimulus durations. Suppose that, after pecking the warning signal, the bird stands in front of the choice key affiliated with the shorter stimulus duration. If the stimulus light offsets before some critical time value is reached, the pigeon will peck that key. If, however, the critical time value is reached before the light offsets, the pigeon moves to the other choice key, thus predisposing the bird to making a correct response *before* the longer time value has even elapsed.

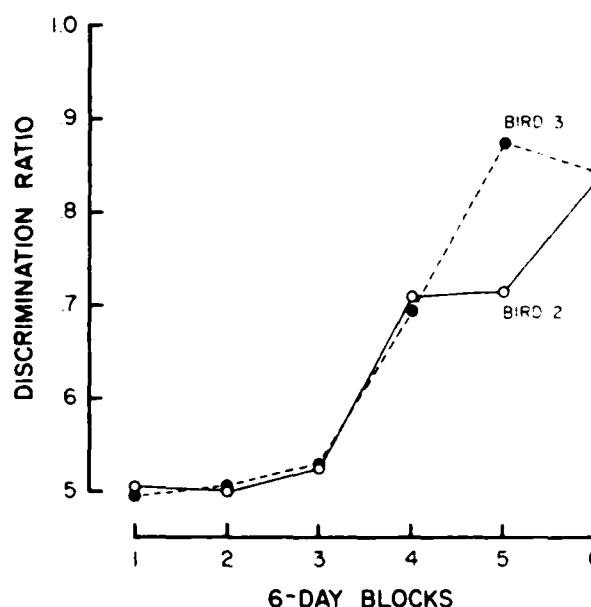
One way to remedy this shortcoming is to present two visually distinctive stimuli on the choice keys and to vary their spatial location from trial to trial. If selection of one test stimulus were correct after the shorter interval and selection of the other test stimulus were correct after the longer interval, then accurate test performance would require the subject to attend to both stimulus durations—in their entirety. This technique was employed by Stubbs¹⁷; it supported orderly discriminative performance by pigeons.

Our own¹⁸ method of separating temporal discrimination from the performance

TABLE 3. Procedure Used in the Study of Duration Discrimination

Trial Type	Warning Stimulus (FR 1)	Duration Stimulus (red)	Trial Sequence		
			Test Stimulus (5.0 sec)	Trial Outcome (3.0 sec)	Intertrial Interval (20.0 sec)
1	White	2 sec	Vertical	Food	
2	White	2 sec	Slanted	Blackout	
3	White	16 sec	Vertical	Blackout	
4	White	16 sec	Slanted	Food	

FIGURE 4. Overall temporal discrimination ratios as a function of 6-day blocks of training. Discrimination ratios were calculated by dividing the rate of response on reinforced trials by the combined rate of response on reinforced and nonreinforced trials.



test used the task shown in TABLE 3. Here, pecking responses to a vertical line test stimulus were reinforced after 2.0 sec of red key illumination, but not to a line test stimulus slanted 60 degrees from vertical; conversely, responses to the slanted-line test stimulus were reinforced after 16.0 sec of red key illumination, but not to a vertical line test stimulus. Here, too, discriminative test responding would have to be based upon the subject's discriminating the *full* durations of *both* temporal stimuli.

Discrimination

Two birds were given 36 days of training using the method outlined in TABLE 3. Each session comprised a total of 40 trials, 10 of each type presented in random blocks. From the rates of responding on reinforced and nonreinforced test trials, an overall discrimination ratio was computed that divided the rate of response on reinforced trials by the combined rate of response on reinforced and nonreinforced trials. These ratios are depicted over 6-day blocks in FIGURE 4. It is readily apparent that discriminative performance rose from a chance level of 0.5, ultimately exceeding a ratio of 0.8 for both birds. As in our earlier work on the temporal-order problem, discrimination learning was disclosed by a decrease in erroneous responding on nonreinforced trials.

In order to get a fuller picture of the pigeons' discriminative capabilities, the initial temporal problem was expanded to include a range of short durations (2.0, 4.0, 6.0, and 8.0 sec) and a range of long durations (10.0, 12.0, 14.0, and 16.0 sec).¹⁷ Reinforcement was available on short-vertical and long-slanted trials; reinforcement was not available on short-slanted and long-vertical trials. Training lasted 96 days, with each daily session comprising a total of 80 trials, 5 of each type (8 durations $\times$ 2 tests) presented in random blocks.

FIGURE 5 shows discriminative performance portrayed as a function of the duration of the temporal stimulus over the final 24 days of training. Discrimination ratios again compared responding on reinforced trials to responding on both reinforced and nonreinforced trials, but here at each of the eight stimulus durations. Discrimination ratio was a direct function of distance from the reinforcement cutoff of 9 sec

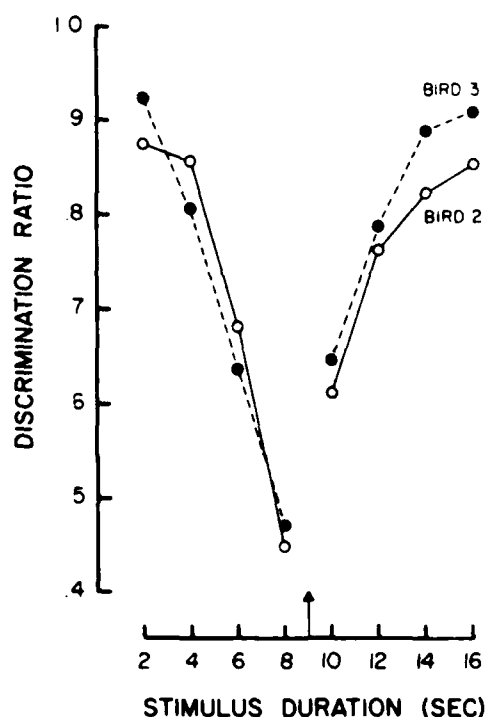


FIGURE 5. Temporal discrimination ratios as a function of the duration of the temporal stimulus. The duration cutoff was 9 sec.

Increasingly different time values were thus easier for pigeons to discriminate than were increasingly similar time values. Additionally, each bird also showed a trend for the short segment of the discriminability function to begin at a lower level and to rise more rapidly than did the long segment. This result is consistent with the birds' overestimating intermediate durations of the key light stimulus: behaving in accord with go/no-go dispositions appropriate to durations a bit longer than those actually given.

Retention

Having determined with our conditional discrimination procedure that pigeons discriminate the durations of key light stimuli and do so with increasing accuracy the more dissimilar the time values, we were interested in ascertaining the pigeons' short-term memory of stimulus durations. To accomplish that goal, we simply inserted retention intervals between offset of the red key light and onset of the vertical or slanted-line test stimulus, in a within-sessions design. If pigeons do forget the prior duration of the red stimulus, we would expect discrimination ratios to fall as the retention interval is lengthened.

Birds 2 and 3 continued as subjects. Prior to the phases of training depicted in FIGURE 6, the birds received extensive training with a new set of duration values: the short stimuli were 1.0, 2.0, 3.0, and 4.0 sec and the long stimuli were 5.0, 6.0, 7.0, and 8.0 sec. Reinforcement was available only on short-vertical and long-slanted trials; the other trial types entailed nonreinforcement. In addition, bird 2 received training with 1.0-, 2.0-, and 4.0-sec retention intervals, and bird 3 received training with 1.0-, 2.0-, 4.0-, and 8.0-sec retention intervals (see DeLong¹⁸ for details).

During the phases depicted in FIGURE 6, birds 2 and 3 received 120 days of

training. For bird 2, daily sessions involved 48 trials, 1 of each possible combination (3 retention intervals $\times$ 8 durations $\times$ 2 tests). For bird 3, daily sessions involved 64 trials, 1 of each possible combination (4 retention intervals $\times$ 8 durations $\times$ 2 tests).

Discriminative performance from the last 24 days in each of these phases is shown in FIGURE 6. Although the overall performance of bird 2 was poorer than that of bird 3, the collective pattern of results was quite similar. As in FIGURE 5, discrimination improved as more extreme stimulus durations were given. Also as in FIGURE 5, the short segment of the discriminability function started at a lower level and was steeper than the long segment. Finally, and most germane to the issue of short-term memory, discrimination was a strong negative function of the retention interval separating the duration to be discriminated and the subsequent performance test. Thus, like other sensory information, the duration of a stimulus becomes increasingly unlikely to control discriminative performance the longer it has been since stimulus offset.¹⁸ Yet even at a retention interval of 16.0 sec, we found measurable control over behavior by stimulus duration.

A final issue to be considered here concerns the way that increases in the retention interval might lead to a worsening of control by prior stimulus duration.¹⁹ One means involves the subject's forgetting the duration of the red key light as the time since stimulus *offset* and the performance test is lengthened. Another possibility is that subjects are basing their test responding on the passage of time since stimulus *onset*; lengthening the retention interval ought then to lead to poorer discrimination and responding more appropriate to long stimulus durations. A third possibility is that subjective estimates of stimulus duration decrease as the retention interval is increased²⁰; subjective shortening will then lower discriminative performance and support responding more appropriate to short stimulus durations. Scrutiny of FIGURE 6 reveals that lengthening the retention interval reduced discrimination scores, but did

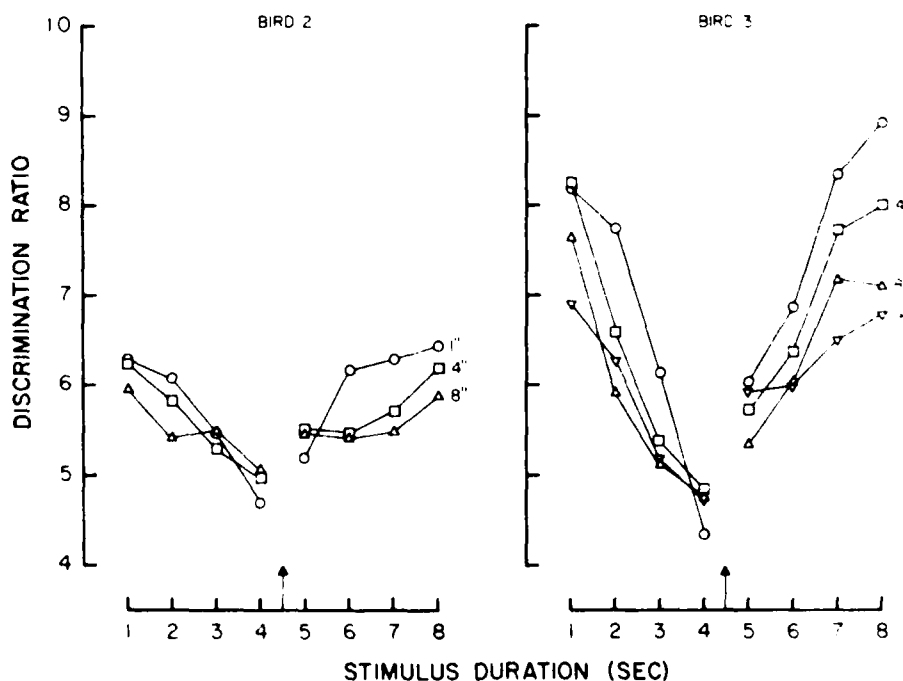


FIGURE 6. Temporal discrimination ratios as functions of the duration of the temporal stimulus and the retention interval. The duration cutoff was 4.5 sec.

not bias performance in any systematic manner. Thus, the latter two interpretations seem not to have been supported by our temporal discrimination data.

CONCLUSIONS

Conditional discrimination procedures reveal that temporal order and duration can serve as discriminative stimuli for pigeon subjects. Not only do pigeons respond discriminatively after different orders of two nonidentical stimuli and different durations of a single stimulus, but they forget this information if the performance test is delayed for several seconds. Discrimination and retention of temporal order and duration suggest that pigeons are able to abstract nonvisual and relational properties from simple visual stimuli.

REFERENCES

1. GIBSON, E. J. 1969. *Principles of Perceptual Learning and Development*. Appleton-Century-Crofts, New York, NY.
2. MARKS, L. E. 1978. *The Unity of the Senses*. Academic Press, New York, NY.
3. SPETCH, M. L. & D. M. WILKIE. 1981. Duration discrimination is better with food access as the signal than with light as the signal. *Learn. Motiv.* **12**: 40-64.
4. LAREW, M. B. 1979. *Discrimination and Retention of Stimulus Order by Pigeons*. Honors thesis, University of Iowa, Iowa City, IA.
5. WASSERMAN, E. A., K. R. NELSON & M. B. LAREW. 1980. Memory for sequences of stimuli and responses. *J. Exp. Anal. Behav.* **34**: 49-59.
6. WEISMAN, R. G., E. A. WASSERMAN, P. W. D. DODD & M. B. LAREW. 1980. Representation and retention of two-event sequences in pigeons. *J. Exp. Psychol. Animal Behav. Processes* **6**: 312-325.
7. WASSERMAN, E. A., S. R. FRANKLIN & E. HEARST. 1974. Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *J. Comp. Physiol. Psychol.* **86**: 616-627.
8. DELONG, R. E. & E. A. WASSERMAN. 1981. Effects of differential reinforcement expectancies on successive matching-to-sample performance in pigeons. *J. Exp. Psychol. Animal Behav. Processes* **7**: 394-412.
9. KONORSKI, J. A. 1959. A new method of physiological investigation of recent memory in animals. *Bull. Acad. Polon. Sci. Ser. Biol.* **7**: 115-117.
10. NELSON, K. R. & E. A. WASSERMAN. 1981. Stimulus asymmetry in the pigeon's successive matching-to-sample performance. *Bull. Psychon. Soc.* **18**: 343-346.
11. WASSERMAN, E. A. 1976. Successive matching-to-sample in the pigeon: Variations on a theme by Konorski. *Behav. Res. Methods Instrum.* **8**: 278-282.
12. BLOUGH, D. S. 1959. Delayed matching in the pigeon. *J. Exp. Anal. Behav.* **2**: 151-160.
13. NELSON, K. R. & E. A. WASSERMAN. 1978. Temporal factors influencing the pigeon's successive matching-to-sample performance: Sample duration, intertrial interval, and retention interval. *J. Exp. Anal. Behav.* **30**: 153-162.
14. HONIG, W. K. & E. A. WASSERMAN. 1981. Performance of pigeons on delayed simple and conditional discriminations under equivalent training procedures. *Learn. Motiv.* **12**: 149-170.
15. FERSTER, C. B. & B. F. SKINNER. 1957. *Schedules of Reinforcement*. Appleton-Century-Crofts, New York, NY.
16. CATANIA, A. C. 1970. Reinforcement schedules and psychophysical judgments: A study of some temporal properties of behavior. *In The Theory of Reinforcement Schedules*. W. N. Schoenfeld, Ed. Appleton-Century-Crofts, New York, NY.
17. STUBBS, A. 1968. The discrimination of stimulus duration by pigeons. *J. Exp. Anal. Behav.* **11**: 223-238.

18. DELONG, R. E. 1983. Control of Responding by Stimulus Duration. Ph.D. dissertation, University of Iowa, Iowa City, IA.
19. CHURCH, R. M. 1980. Short-term memory for time intervals. *Learn. Motiv.* **11**: 208-219.
20. SPETCH, M. L. & D. M. WILKIE. 1983. Subjective shortening: A model of pigeons' memory for event duration. *J. Exp. Psychol. Anim. Behav. Processes* **9**: 14-30.

Contingent Aftereffects in Duration Judgments^a

LORRAINE G. ALLAN

*Department of Psychology
McMaster University
Hamilton, Ontario, Canada L8S 4K1*

Exposure to a repeating event of a constant duration has been shown to influence the judged duration of subsequently presented test events. For example, adaptation to a 4-sec tone results in the shortening of the judged duration of a subsequently present 2-sec test tone.¹ This duration aftereffect is negative in that the judged duration of the test tone is in a direction away from the value of the adaptation duration. Recently, negative duration aftereffects which are contingent on the pitch of the tone or on the temporal order of presentation of two events have been reported.²⁻⁴ These contingent duration aftereffects are reminiscent of the contingent aftereffect discovered by McCollough.⁴ Her subjects inspected two visual patterns that alternated every few seconds. Under one condition, one adaptation figure consisted of black vertical bars on an orange background and the other adaptation figure consisted of black horizontal bars on a blue background. During the test, the subject viewed achromatic patterns and reported that the achromatic background of the vertical bars appeared bluish and the achromatic background of the horizontal bars appeared orangeish. McCollough⁴ argued that this orientation-contingent negative aftereffect could be understood in terms of color adaptation of orientation-specific edge detectors and she interpreted her result as psychophysical evidence for neural units that are both color- and orientation-specific. While selective adaptation of neural detectors is still the favored explanation,⁵ there are data that are supportive of an associative account.^{6,7}

In the reports^{2,3} of contingent duration aftereffects an attempt was made to explain the data using the two accounts that have been applied to the McCollough effect. Both selective adaptation of neural detectors and associative mechanisms are far removed from the usual theoretical concepts found in the time perception literature. The research to be reported questions the assumption that the contingent duration aftereffects are related to the McCollough effect.

There is considerable evidence in the time perception literature of systematic "biases" in judgments of perceived duration.⁸ Duration judgements are influenced by such nontemporal characteristics of the duration marker as energy, complexity, modality, and whether the interval is filled or empty. It is possible that the contingent duration aftereffects are yet another example of judged duration's being influenced by variables in addition to physical duration.

The experiments to be reported compare contingent duration aftereffects with the McCollough effect. A number of studies have shown that McCollough-type contingent aftereffects last for days and even weeks.⁹ One purpose was to study the persistence of the contingent duration aftereffects. In one experiment a delay was introduced between adaptation and test and its effect on the size of the order-contingent duration

^aThis research was supported by the Natural Sciences and Engineering Research Council of Canada. Some of the data reported in Experiments 3 and 4 were presented at the 1980 meeting of the Psychonomic Society.

aftereffect was examined. In addition, some subjects ran multiple sessions under both adaptation conditions so that the persistence of the aftereffects over sessions could be studied and the aftereffects for individual subjects could be examined.

For McCollough-type contingent aftereffects, whenever the perception of feature *A* has been found to be contingent on feature *B*, the perception of *B* has also been shown to be contingent on *A*. For example, judgment of color is contingent on line orientation, on spatial frequency, and on motion direction; and judgment of line orientation, spatial frequency, and motion direction is contingent on color.⁵ If judgments of duration can be contingent on pitch, one might expect judgments of pitch to be contingent on duration. Another purpose was to determine whether a negative pitch aftereffect contingent on duration could be demonstrated.

A magnitude estimation task² and a reproduction task³ have been used in the studies of contingent duration aftereffects. To extend the generality of results, another psychophysical procedure, forced-choice discrimination, is tried.

PROCEDURE

One hundred twenty-five undergraduates at McMaster University participated as part of a course requirement. In addition, data were obtained from six paid subjects who ran multiple sessions.

Stimulus presentation, recording of responses, and timing were controlled either by a PDP 8/E computer with a video terminal or by a SuperPET. Auditory signals were presented binaurally over headphones. A Wavetek Function Generator was used with the PDP 8/E to produce a 70-dB pure tone signal. Otherwise, the SuperPET's tone generator was used.

In all experiments a session was made up of a preadaptation phase and three adaptation-test sequences. Preadaptation and test consisted of discrimination trials. A trial began with READY displayed on the screen. This was followed by a pair of tones. To insure that subjects make a comparative judgment on each trial, a roving standard design was used. At the end of the second tone, RESPOND was displayed on the screen. The subject indicated his judgment by typing a 1 or a 2 on the keyboard. During adaptation the subject simply listened to a series of tones.

EXPERIMENT 1: DURATION CONTINGENT ON TEMPORAL ORDER^a

Method

Forty subjects were randomly divided into four groups of 10 subjects each. Preadaptation consisted of 96 duration discrimination trials. READY was displayed for 500 msec and 1 sec later a pair of 364-Hz tones, separated by 400 msec, was presented. The two tones differed in duration by 30 msec. There were four pairs of duration values: 330/360, 360/390, 390/420, and 420/450. Each pair was presented in two orders, short followed by long (SL) or long followed by short (LS). Twenty subjects were asked whether the longer tone occurred first or second ("long instruction") and 20 were asked whether the shorter tone occurred first or second ("short instruction").

^aParts of Experiments 1 and 2 are described in a B.Sc. thesis by A. M. Majerovich¹ at McMaster University.

The subject responded by typing 1 (for first) or 2 (for second) on the computer keyboard. The intertrial interval was 1.4 sec.

During adaptation the subject listened to 50 pairs of tones. The two tones in a pair were separated by 400 msec and pairs were separated by 1.4 sec. One tone in a pair was 200 msec, the other was 600 msec. For 10 subjects under each instruction, the first tone was the short (short-long adaptation); for the remaining 10 subjects under each instruction, the first tone was longer (long-short adaptation). Thirty-two duration discrimination test trials immediately followed adaptation. The adaptation test sequence occurred three times for a total of 96 test trials.

Four subjects ran repeat sessions under similar conditions. The difference in duration was 20 msec and only the long instruction was used. The four pairs of duration values were 370/390, 390/410, 410/430, 430/450. Two subjects (M.G. and H.B.) ran under long-short adaptation and then under short-long; the other two subjects (L.P. and D.K.) received the adaptation conditions in the reverse order. The number of sessions under each adaptation condition is shown in TABLE 2.

Results and Discussion

The probability of a correct duration judgment during preadaptation and during test is present in TABLE 1 for each of the four groups. The data are summarized by the difference between two conditional probabilities, known as the time-order error (TOE):

$$\text{TOE} = P(R_{LS}|LS) - P(R_{SL}|SL)$$

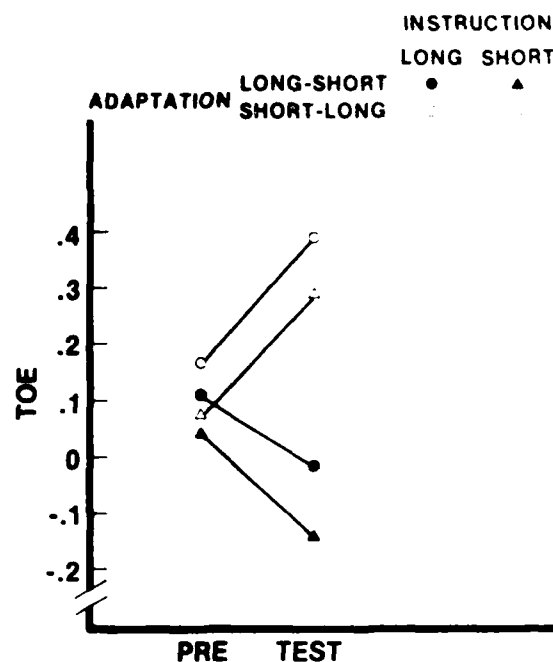
where $P(R_{LS}|LS)$ is the probability of a correct response to the *LS* order and $P(R_{SL}|SL)$ is the probability of a correct response to the *SL* order. Most research has indicated that for brief durations TOE is positive and for longer durations it is negative.⁸ TABLE 1 shows that for all four groups in this experiment, TOE is positive during preadaptation. While the mean size of preadaptation TOE is smaller under the short instruction, instruction is not a significant variable according to the Mann-Whitney *U* test ($U = 159.5$, $p > .05$). This result is consistent with other findings.^{10,11}

If a negative duration aftereffect can be made contingent on temporal order, then after short-long adaptation, the first test tone presented should appear longer compared to preadaptation judgments and the second presented test tone should appear shorter, resulting in a more positive TOE in test than preadaptation. After

TABLE 1. $P(R_{LS}|LS)$, $P(R_{SL}|SL)$ and TOE: Group Data from Experiments 1 and 2

Instruction	Adaptation Condition	Preadaptation			Test		
		$P(R_{LS} LS)$	$P(R_{SL} SL)$	TOE	$P(R_{LS} LS)$	$P(R_{SL} SL)$	TOE
<i>Experiment 1</i>							
Long	Long-short	.73	.62	.11	.69	.71	.02
	Short-long	.76	.59	.17	.81	.42	.39
Short	Long-short	.66	.62	.04	.56	.70	.14
	Short-long	.67	.59	.08	.74	.45	.29
<i>Experiment 2</i>							
Long	Long-short	.69	.54	.15	.73	.57	.16
	Short-long	.76	.63	.13	.79	.66	.13

FIGURE 1. Preadaptation and test TOE in Experiment 1 as a function of adaptation condition. The data are shown separately for each instruction.



long-short adaptation, the first presented test tone should appear shorter compared to preadaptation judgments and the second presented test tone should appear longer, resulting in a less positive TOE during test than during preadaptation. TABLE 1 and FIGURE 1 show these predictions to be supported by the data. For 30 of the individual subjects, the change in TOE was in the expected direction: 19 in the short-long groups and 11 in the long-short groups. A χ^2 test indicates that the proportion of the subjects showing a change in TOE in the expected duration was significant ($\chi^2 = 10.0$, $p < .005$). These data demonstrate that judged duration is influenced by the pair of stimuli presented during adaptation. One adaptation condition increases the size of the preadaptation positive TOE; the other changes the preadaptation positive TOE to a negative one.

$P(R_{LS} | LS)$, $P(R_{SL} | SL)$, and TOE, averaged over all sessions, are shown in TABLE 2 for each of the four repeat-session subjects. In FIGURE 2, preadaptation and test TOE are plotted for each adaptation condition. D.K. is the only subject who shows no effect of adaptation. For the other three subjects, long-short adaptation results in a more negative TOE during test than during preadaptation and short-long adaptation in a more positive TOE. Thus, for an individual subject, TOE can be made more positive or more negative depending upon whether the subject is exposed to short long or long-short pairs of durations.

The overall probability of a correct duration response, regardless of the presentation order, is also shown in TABLE 2, where

$$P(C) = P(LS)P(R_{LS} | LS) + P(SL)P(R_{SL} | SL).$$

For subject H.B., $P(C)$ increases from preadaptation to test under both adaptation conditions, whereas for the other three subjects, $P(C)$ decreases from preadaptation to test under both adaptation conditions. Thus, while adaptation has a differential effect on reports about perceived duration, here measured by TOE, adaptation does not differentially affect discriminability, as measured by $P(C)$. This result is consistent

TABLE 2. $P(R_{LS}|LS)$, $P(R_{SL}|SL)$, $P(C)$ and TOE: Repeat-Session Subject Data from Experiment 1

Adaptation Condition and Subject	Preadaptation				Test			
	$P(R_{LS} LS)$	$P(R_{SL} SL)$	$P(C)$	TOE	$P(R_{LS} LS)$	$P(R_{SL} SL)$	$P(C)$	TOE
<i>Long-Short</i>								
M.G. (8) ^a	.70	.77	.73	.07	.62	.82	.72	.20
H.B. (8) ^a	.64	.55	.60	.09	.57	.66	.61	.09
L.P. (8) ^b	.66	.59	.63	.07	.48	.69	.58	.21
D.K. (7) ^b	.64	.66	.65	-.02	.61	.63	.62	.02
Mean				.02				.13
<i>Short-Long</i>								
M.G. (8)	.80	.71	.76	.09	.85	.55	.70	.30
H.B. (7)	.56	.59	.57	-.03	.63	.60	.61	.03
L.P. (6)	.59	.55	.57	.04	.67	.45	.56	.23
D.K. (10)	.52	.66	.59	-.14	.50	.64	.57	.14
Mean				-.01				.10

NOTE: Number of sessions under each adaptation condition is shown in parentheses.

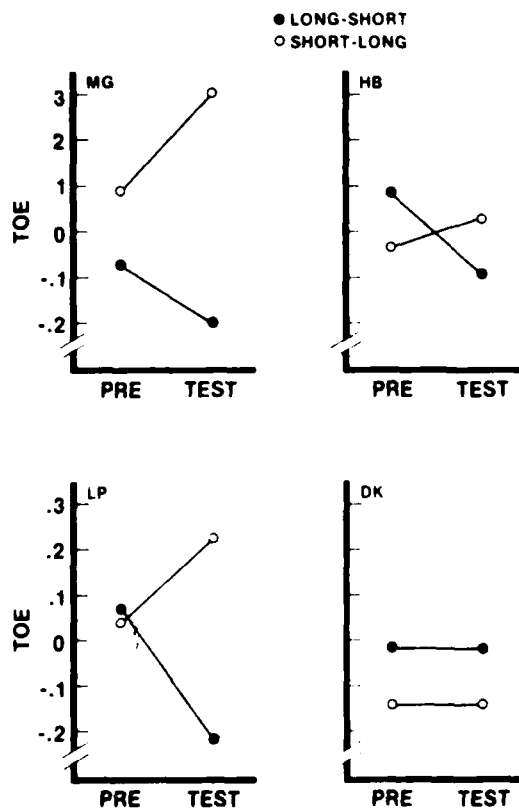
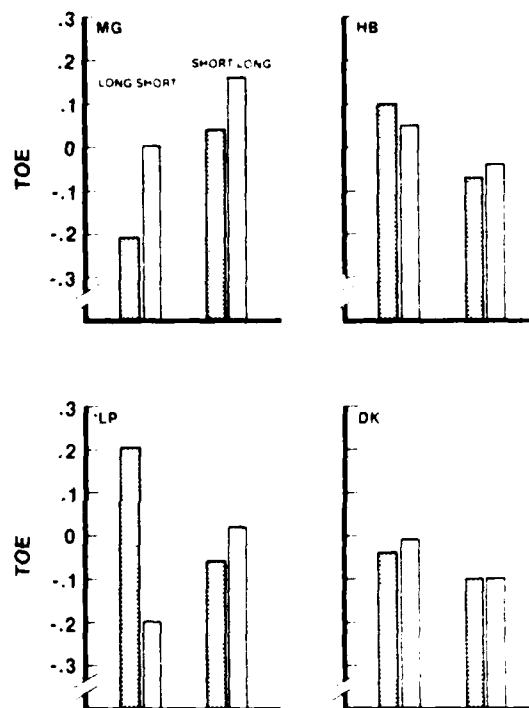
^aReceived the long-short condition first.^bReceived the short-long condition first.

FIGURE 2. Preadaptation and test TOE in Experiment 1 as a function of adaptation condition. The data are shown separately for each repeat-session subject

with other findings in the literature that variables that influence reports about perceived duration often have no differential effect on discriminability.⁷

Preadaptation TOE is shown in FIGURE 3, separately for sessions 1 and 2 and for sessions ($N-1$) and N (the last two sessions). If the order-contingent duration aftereffect showed persistence over sessions, one would expect preadaptation TOE to become more negative over sessions under long-short adaptation and more positive over sessions under short-long adaptation. Subject L.P. shows this pattern. Subjects H.B. and D.K. show little change in preadaptation TOE over sessions, and M.G. shows

FIGURE 3. Preadaptation TOE in Experiment 1 shown separately for early and late sessions. *Stippled bars* represent mean TOE during sessions 1 and 2 and *open bars* during sessions ($N-1$) and N . The pair of bars on the *left* are for the long short condition and the pair on the *right* are for the short-long condition.



the reverse trend under the long-short condition. These data suggest that the adaptation effect does not persist over sessions.

EXPERIMENT 2: PERSISTENCE OF THE ORDER-CONTINGENT DURATION AFTEREFFECT

Method

Twenty subjects were randomly divided into two groups of 10 subjects each. The method was similar to that used in Experiment 1, except that a 25-min delay was introduced between the end of each adaptation period and the beginning of test. The subjects remained in the experimental room and were instructed to keep the headphones on. Only the long instruction was used.

Results and Discussion

The probability of a correct duration judgment during preadaptation and during test and the value of TOE are presented in TABLE 1 for each adaptation group. As in Experiment 1, TOE is positive during preadaptation. Overall, there is no change in TOE from preadaptation to test following either adaptation condition. In the long-short group only five subjects showed the expected change in TOE (more negative) and in the short-long group only four subjects showed the expected change in TOE (more positive). These data suggest that the influence of the adaptation condition has disappeared by the end of a 25-min delay period.

EXPERIMENT 3: DURATION CONTINGENT ON PITCH

Method

Thirty-four subjects were randomly divided into two groups of 17 subjects each. Preadaptation consisted of 96 duration discrimination trials. READY was displayed for 500 msec and 1 sec later a pair of tones, separated by 400 msec, was presented. One tone was 600 Hz, the other 900 Hz. The two tones also differed in duration by 35 msec. There were four pairs of duration values: 330/365, 365/400, 400/435, and 435/470. Two duration orders and two frequency orders resulted in four events, each event occurring 24 times in the 96 trials. The subject was required to indicate which tone, first or second, was longer by typing 1 or 2 on the terminal keyboard. The intertrial interval was 2 sec.

During adaptation the subject listened to two alternating tones, with a 400-msec interval between successive tones. Each tone was presented 100 times. For the *Lo/L-Hi/S* group, one tone was presented at 600 Hz for 600 msec and the other at 900 Hz for 200 msec; for the *Lo/S-Hi/L* group, one tone was presented at 600 Hz for 200 msec and the other at 900 Hz for 600 msec. Sixteen discrimination test trials immediately followed adaptation. The adaptation-test sequence occurred three times for a total of 48 test trials.

Six subjects ran repeat sessions under similar conditions. The difference in duration between the two tones was 20 msec and the four pairs of duration values were 370/390, 390/410, 410/430, 430/450. The intertone interval and the intertrial interval were 1 sec. During adaptation each tone was presented 50 times with a 1-sec interval between tones. Thirty-two discrimination trials followed adaptation. The three adaptation-test sequences resulted in 96 test trials. Three subjects (K.C., D.K., and I.S.) ran under *Lo/L-Hi/S* adaptation and then under *Lo/S-Hi/L*; the other three subjects (H.B., M.G., and L.P.) received the adaptation conditions in the reverse order. The number of sessions under each adaptation condition is shown in TABLE 4.

Results and Discussion

The probability of a correct duration judgment during preadaptation and during test is presented in TABLE 3 for each adaptation group. The data are organized in matrices where the rows represent duration order, short followed by long (SL) or long followed by short (LS), and the columns represent frequency order, low followed by high (LH) or high followed by low (HL). The negative diagonal of the preadaptation matrices was larger than the positive for 22 of the 34 subjects and was the same for 4

subjects. The Wilcoxon T test shows the sum of the entries along the negative diagonal to be significantly larger than the sum along the positive diagonal ($T = 116.5$, $n = 30$, $p < .02$). These data indicate that short low-pitched tones are judged shorter than short high-pitched tones and that long high-pitched tones are judged longer than long low-pitched tones. This dependence of duration judgments on frequency is referred to as auditory *kappa* in the literature.^{13,14} In the remainder of this paper *kappa* will be defined as:

$$[P(R_{SL}|SL \text{ and } LH) + P(R_{LS}|LS \text{ and } HL)] \\ - [P(R_{SL}|SL \text{ and } HL) + P(R_{LS}|LS \text{ and } LH)]$$

If a negative duration aftereffect can be made contingent on pitch, then after *Lo/S-Hi/L* adaptation, low-pitched tones should appear shorter and high-pitched tones should appear longer during test than during preadaptation. That is, *kappa* should be larger after adaptation. The data in TABLE 3 show this to be the case in that the difference between the negative and positive diagonals is greater during test (.43) than during preadaptation (.17). For the *Lo/S-Hi/L* group, low-pitched tones should

TABLE 3. Duration Discrimination Data from Experiment 3

Adaptation Condition	Preadaptation			Test		
	Duration Order	Pitch Order		Duration Order	Pitch Order	
		LH	HL		LH	HL
<i>Lo/L-Hi/S</i>	SL	.74	.63	SL	.80	.58
	LS	.66	.72	LS	.53	.74
<i>Lo/S-Hi/L</i>	SL	.72	.64	SL	.60	.62
	LS	.61	.71	LS	.77	.78

NOTE: Each entry is the probability of a correct duration response.

appear longer and high-pitched tones should appear shorter during test than during preadaptation. That is, *kappa* should be attenuated. The data in TABLE 3 show this to be the case in that the difference between the two diagonals during preadaptation (.18) is eliminated during test (-.01). Of the 34 subjects, 22 showed performance during test, relative to preadaptation, that was in the expected direction (12 with larger *kappa* after *Lo/L-Hi/S* and 10 with reduced *kappa* after *Lo/S-Hi/L*) and one subject showed no change. According to the Wilcoxon T test these data indicate that the adaptation conditions had a significant differential effect on performance ($T = 142.5$, $n = 33$, $p < .02$). One adaptation condition increases the preadaptation tendency of judging low-pitched tones as relatively short and high-pitched tones as relatively long; the other adaptation condition eliminates this tendency.

The probability of a correct duration judgment, averaged over all sessions under a particular adaptation condition, and the value of *kappa* are shown in TABLE 4 for each of the six repeat-session subjects. For five subjects *kappa* was increased after *Lo/L-Hi/S* adaptation and for five subjects *kappa* was decreased after *Lo/S-Hi/L* adaptation. The data averaged over the six subjects show the same patterns as do the group data in TABLE 3. There is a *kappa* effect during preadaptation. *Kappa* is

increased from .18 to .30 by *Lo/L Hi/S* and *kappa* is decreased from .12 to .06 by *Lo/S-Hi/L*.

The overall probability of a correct response, regardless of presentation order and pitch, is shown in TABLE 5 for each of the six repeat-session subjects. Adaptation condition does not have a systematic effect on $P(C)$ across subjects. As in Experiment 1, while adaptation condition has a differential effect on reports about perceived duration, here measured by *kappa*, adaptation condition does not differentially influence discriminability, as measured by $P(C)$. Again, these results are consistent with other findings that variables that affect reports about perceived duration often do not influence discriminability.⁸

Preadaptation *kappa* is shown in FIGURE 4, separately for sessions 1 and 2 and sessions ($N-1$) and N . If the pitch-contingent duration aftereffect showed persistence

TABLE 4. Probability of a Correct Duration Response for Each of the Four Stimulus Events and the Value of Kappa for the Six Repeat-Session Subjects in Experiment 3

Adaptation Condition and Subject	Preadaptation					Test				
	<i>LH</i>		<i>HL</i>		Kappa	<i>LH</i>		<i>HL</i>		Kappa
	<i>SL</i>	<i>LS</i>	<i>SL</i>	<i>LS</i>		<i>SL</i>	<i>LS</i>	<i>SL</i>	<i>LS</i>	
<i>Lo/L-Hi/S</i>										
H.B. (6) ^a	.77	.39	.73	.55	.20	.76	.44	.61	.62	.32
L.P. (8) ^a	.63	.56	.45	.69	.31	.59	.51	.34	.78	.52
M.G. (8) ^a	.63	.81	.76	.69	.26	.60	.81	.68	.80	.08
K.C. (7) ^b	.70	.37	.78	.42	.03	.70	.26	.62	.49	.31
D.K. (7) ^b	.87	.54	.68	.76	.41	.81	.48	.63	.76	.47
I.S. (7) ^b	.74	.43	.53	.67	.45	.70	.39	.67	.63	.26
Mean					.18					.30
<i>Lo/S-Hi/L</i>										
H.B. (8)	.64	.53	.65	.52	-.02	.60	.58	.67	.52	.14
L.P. (9)	.68	.56	.51	.71	.32	.57	.67	.55	.72	.07
M.G. (7)	.51	.90	.74	.65	.48	.46	.86	.73	.69	.45
K.C. (8)	.70	.44	.56	.65	.35	.61	.44	.57	.57	.17
D.K. (6)	.82	.51	.65	.75	.41	.76	.73	.79	.65	.10
I.S. (7)	.82	.46	.73	.52	.14	.76	.51	.70	.55	.10
Mean					.12					.06

NOTE: Number of sessions under each adaptation condition is shown in parentheses.

^aReceived the *Lo/S Hi/L* condition first.

^bReceived the *Lo/L Hi/S* condition first.

over sessions, one would expect preadaptation *kappa* to increase over sessions under *Lo/L Hi/S* and to decrease over sessions under *Lo/S Hi/L*. A consistent pattern across subjects is not seen. These data, like the data from the earlier experiments, suggest that the influence of adaptation does not persist over sessions.

EXPERIMENT 4: PITCH CONTINGENT ON DURATION

Method

Thirty-one subjects were randomly divided into two groups of 16 subjects and 15 subjects. Preadaptation consisted of 96 pitch discrimination trials. READY was

TABLE 5. Overall Probability of a Correct Duration Response, $P(C)$, in Experiment 3

Condition and Subject	$P(C)$	
	Preadaptation	Test
<i>Lo/L-Hi/S</i>		
H.B.	.61	.61
L.P.	.58	.56
M.G.	.72	.72
K.C.	.57	.52
D.K.	.71	.67
I.S.	.59	.59
Mean	.63	.61
<i>Lo/S-Hi/L</i>		
H.B.	.58	.59
L.P.	.62	.63
M.G.	.70	.68
K.C.	.59	.55
D.K.	.68	.74
I.S.	.63	.63
Mean	.63	.64

displayed for 500 msec and 1 sec later a pair of tones, separated by 400 msec, was presented. The two tones differed in duration (200 msec or 600 msec). The two tones also differed in frequency by 8 Hz. There were four pairs of frequency values: 732/740, 740/748, 748/756, and 756/764. The subject was required to indicate which tone, first or second, was higher in pitch by typing 1 or 2 on the terminal keyboard.

The adaptation conditions were identical to those in Experiment 3. The subject listened to two alternating tones, with a 400-msec interval between successive tones. Each tone was presented 100 times. There were 16 subjects in the *Lo/L-Hi/S* group

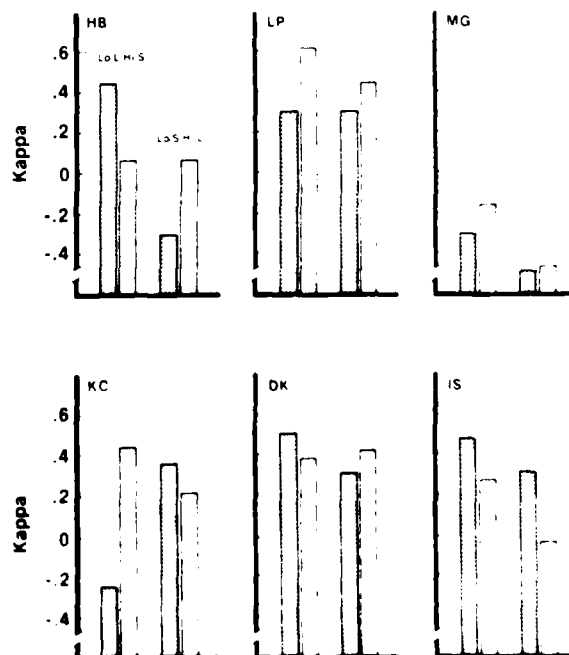


FIGURE 4. Preadaptation kappa in Experiment 3 shown separately for early and late sessions. *Stippled bars* represent mean kappa during sessions 1 and 2 and *open bars* during sessions (N-1) and N. The pair of bars on the left are for the *Lo/L-Hi/S* condition and the pair on the right are for the *Lo/S-Hi/L* condition.

and 15 in the *Lo/S-Hi/L* group. Sixteen pitch discrimination test trials immediately followed each adaptation period. The adaptation-test sequence occurred three times for a total of 48 test trials.

Six subjects ran repeat sessions under similar conditions. The difference in pitch between the two tones was 9 Hz and the four pitch pairs were 726/735, 735/744, 744/753, 753/762. The intertone interval and intertrial interval were 1 sec. During adaptation each tone was presented 50 times with a 1-sec interval between tones. Thirty-two discrimination trials followed adaptation. The three adaptation test sequences resulted in 96 test trials. Three subjects (D.K., L.P., and K.C.) ran under *Lo/L-Hi/S* adaptation and then under *Lo/S-Hi/L*; the other three subjects (H.B., M.G., and I.S.) received the adaptation conditions in the reverse order. The number of sessions under each adaptation condition is shown in TABLE 7.

Results and Discussion

The data in TABLE 6 are presented in a similar manner to those in TABLE 3, except that the entries represent the probability of a correct pitch judgment. The negative diagonal of the preadaptation matrices was larger than the positive for 21 of the 31 subjects. The Wilcoxon T test shows the sum of the entries along the negative diagonal to be significantly larger than the sum along the positive diagonal ($T = 103.5$, $n = 31$, $p < .01$). These data show that low-frequency short tones are judged lower in pitch than low-frequency long tones and that high-frequency long tones are judged higher in pitch than high-frequency short tones. The dependence of pitch judgments on duration is referred to as auditory *tau* in the literature.^{13,15} In the remainder of this paper *tau* will be defined as:

$$[P(R_{HL}|LH \text{ and } SL) + P(R_{HL}|HL \text{ and } LS)] \\ - [P(R_{LH}|LH \text{ and } LS) + P(R_{HL}|HL \text{ and } SL)]$$

For the *Lo/L-Hi/S* group, short tones should appear lower in pitch and long tones higher in pitch during test than during preadaptation. That is, *tau* should be larger after adaptation. The data in TABLE 6 indicate that there is little change in *tau* from preadaptation (.32) to test (.29). If anything, *tau* is reduced. For the *Lo/S-Hi/L* group, short tones should appear higher in pitch and long tones lower in pitch during test than during preadaptation. That is, *tau* should be attenuated. The data in TABLE 6 show this to be the case in that the difference between the negative and positive

TABLE 6. Pitch Discrimination Data from Experiment 4

Adaptation Condition	Preadaptation		Test	
	Duration Order	Pitch Order <i>LH</i> <i>HL</i>	Duration Order	Pitch Order <i>LH</i> <i>HL</i>
<i>Lo/L-Hi/S</i>	<i>SL</i>	.82 .66	<i>SL</i>	.89 .67
	<i>LS</i>	.66 .82	<i>LS</i>	.77 .84
<i>Lo/S-Hi/L</i>	<i>SL</i>	.86 .60	<i>SL</i>	.81 .70
	<i>LS</i>	.64 .77	<i>LS</i>	.68 .81

NOTE: Each entry is the probability of a correct pitch response.

TABLE 7. Probability of a Correct Pitch Response for Each of the Four Stimulus Events and the Value of Tau for Each of the Repeat-Session Subjects in Experiment 4

Condition and Subject	Preadaptation					Test				Tau
	SL		LS		Tau	SL		LS		
	LH	HL	LH	HL		LH	HL	LH	HL	
<i>Lo/L-Hi/S</i>										
H.B. (8) ^a	.69	.38	.59	.51	.22	.68	.40	.61	.42	.09
L.P. (9) ^b	.81	.63	.82	.58	-.06	.87	.60	.82	.56	.00
M.G. (8) ^a	.97	.88	.94	.96	.12	.99	.78	.93	.95	.23
K.C. (7) ^b	.83	.35	.54	.56	.50	.79	.35	.51	.48	.42
D.K. (8) ^b	.79	.79	.70	.79	.09	.81	.69	.68	.77	.21
I.S. (8) ^a	.96	.89	.99	.96	.05	.99	.85	1.00	.94	.07
Mean					.15					.17
<i>Lo/S-Hi/L</i>										
H.B. (8)	.65	.50	.65	.53	.02	.68	.35	.61	.46	.18
L.P. (10)	.71	.84	.78	.74	-.17	.60	.85	.70	.78	.16
M.G. (8)	.95	.90	.91	.93	.06	.96	.84	.95	.94	.10
K.C. (6)	.73	.42	.57	.52	.26	.78	.37	.58	.43	.26
D.K. (8)	.84	.87	.85	.81	-.07	.79	.84	.90	.78	.17
I.S. (8)	.81	.84	.88	.90	-.01	.83	.88	.86	.91	.01
Mean					.02					.03

NOTE: The number of sessions under each adaptation condition is shown in parentheses

^aReceived the *Lo/S-Hi/L* condition first.^bReceived the *Lo/L-Hi/S* condition first.

diagonals is less during test (.24) than during preadaptation (.39). Of the 31 subjects, 19 showed performance during test, relative to preadaptation, that was in the expected direction (7 with larger *tau* after *Lo/L-Hi/S* and 12 with reduced *tau* after *Lo/S-Hi/L*), and one showed no change. According to the Wilcoxon T test these data indicate that the adaptation conditions did not have a significant differential effect on performance ($T = 200$, $n = 30$, $p > .05$). Judged pitch does not appear to be differentially influenced by the pair of stimuli presented during adaptation. Both adaptation conditions tended to reduce *tau*.

The probability of a correct pitch judgment, averaged over all sessions under a particular adaptation condition, and the value of *tau* are shown in TABLE 7 for each of the six repeat-session subjects. Only one subject (D.K.) showed the expected result: *tau* increased after *Lo/L-Hi/S* and decreased after *Lo/S-Hi/L*. For the other subjects adaptation did not have a consistent effect. Averaged over the six subjects there is virtually no change in *tau* under either condition. The data from these subjects are in agreement with the group data in TABLE 6. In neither case were we able to establish a negative pitch aftereffect contingent on duration.

DISCUSSION

The data show that judged duration lengthens with increases in pitch and judged pitch increases with duration. While others have reported demonstrations of auditory *tau* and of auditory *kappa*, the effects in the present experiments are clearer and more

consistent. The data also provide strong support for the claim that TOE for relatively brief durations is positive.^{8,16}

The adaptation conditions had a differential effect on TOE in Experiment 1 and on κ in Experiment 3. This can be taken as evidence for order-contingent and pitch-contingent duration aftereffects respectively. In neither of these experiments did the adaptation conditions differentially influence $P(C)$. Thus, adaptation influences reports about perceived duration, but not discriminability.

If duration aftereffects are variants of the McCollough effect, one would expect to be able to demonstrate a duration-contingent pitch aftereffect. We were unable to do so in Experiment 4. The data from the repeat-session subjects suggest that neither the order-contingent nor the pitch-contingent duration aftereffect persists over sessions. Furthermore, the order-contingent duration aftereffect was not observed when a relatively short delay of 25 min was introduced after adaptation. In contrast, the McCollough effect and its variants persist for days, even weeks. Our data suggest fundamental differences between the duration aftereffects and the McCollough effect.

The favored account of the McCollough effect is the selective adaptation of feature detectors. For duration aftereffects, this account would imply duration detectors selectively tuned to temporal order or pitch. We know of no physiological evidence for such detectors.

TOE is a well-documented finding in the time perception literature. A satisfactory account of the phenomenon does not exist. Woodrow¹⁶ suggested that the perceptual duration of the first presented event "gravitated" towards a remote standard, the indifference interval. This would result in a positive TOE for durations less than the indifference interval and a negative TOE for durations greater than the indifference interval. Aside from defining the indifference interval, this account has difficulty with the finding that TOE, whether positive or negative, tends toward zero as the interstimulus interval is increased.^{17,18}

Hellström^{11,12} has presented a model that postulates that the two internal durations on a trial are modified by the adaptation level or mean subjective duration, and the differential weights are then applied to the modified internal values. The response is based on a comparison of these weighted internal values with a criterion internal value. The differential weights associated with the first and second duration values are identified with retroactive and proactive interference effects respectively. This interpretation of the weight parameters is consistent with the finding that TOE tends toward zero as interstimulus interval is increased. Interference effects should decrease with interstimulus interval and therefore so should TOE. In order to account for a positive TOE with brief durations and a negative TOE with longer durations, Hellström has to assume that proactive interference is the greater for short durations and retroactive is the greater for longer durations. Why this should be the case is not developed in the model.

Allan⁹ presented a model for TOE that provided a qualitative account of her data. However, Jamieson²⁰ showed that the model provided a poor quantitative fit to the data.

The data from the present experiments do not account for the finding that the sign of TOE depends upon duration. However, they do provide information about the differential effect of the two adaptation conditions on the preadaptation positive TOE. As was noted above, a positive TOE, as found during preadaptation in the present studies, indicates that R_{TS} responses are made more frequently than R_{ST} responses. The effect of adaptation on this distribution of responses can be explained as contrast in perceived duration between test and adaptation. The difference between the perceived durations of the two tones during long short adaptation is large relative to the test

difference. Therefore, only when the subject is certain during test that the first stimulus is the longer would the trial be labeled R_{LS} . This would result in fewer R_{LS} responses during test than during preadaptation and a more negative TOE. The reverse is the case for short-long adaptation. Again, the difference between the perceived durations of the two tones during adaptation is noticeable relative to the test difference. Now only when the subject is certain during test that the first stimulus is the shorter would the trial be labeled R_{SL} . This would result in more R_{LS} responses during test than during preadaptation and a more positive TOE. In decision theory language, the effect of adaptation is on the criterion. The criterion value increases after long short adaptation and decreases after short-long adaptation.

Kappa and *tau* describe the finding that high-frequency tones and long tones "belong" together as do low-frequency tones and short tones. Thus, when tones are presented in the HL order, R_{LS} is more likely than R_{SL} ; and when tones are presented in the LH order, R_{SL} is more likely than R_{LS} . This is the *kappa* effect. Similarly, R_{HH} is more likely than R_{HH} when tones are presented in the LS order; and R_{LL} is more likely than R_{LL} when tones are presented in the SL order. This is *tau*.

The contrast hypothesis, used to explain the differential effects of adaptation on TOE, can be applied to the differential effects of adaptation on *kappa* as well. During preadaptation, the criterion on HL trials is lower than the criterion on LH trials, resulting in more R_{LS} responses on HL trials than on LH trials and a *kappa* effect. Under the *Lo/S-Hi/L* adaptation condition, the high tone is always longer than the low tone and the difference between the perceived durations of the two tones is large relative to the test difference. Therefore, on HL test trials, only when the subject is certain that the first tone is the longer is the trial labeled R_{LS} , resulting in a decrease in R_{LS} on HL test trials relative to preadaptation; on LH test trials, only when the subject is certain that the first tone is the shorter is the trial labeled R_{SL} , resulting in a decrease in R_{SL} on LH test trials relative to preadaptation. That is, the criterion for R_{LS} is increased on HL trials and is lowered on LH trials, resulting in a reduced *kappa*. Now consider the *Lo/L-Hi/S* adaptation condition. The high tone is always shorter than the low tone. On HL test trials, only when the subject is certain that the first tone is the shorter is the trial labeled R_{SL} , resulting in a decrease in R_{SL} on HL test trials; on LH test trials, only when the subject is certain that the first tone is the longer is the trial labeled R_{LS} , resulting in a decrease in R_{LS} on LH test trials. That is, the criterion for R_{LS} is lowered on HL trials and is increased on LH trials, resulting in an increased *kappa*.

The contrast hypothesis explains the effect of adaptation through criterion shifts and therefore predicts that that discriminability will not be influenced by adaptation. This is what was found. One would not expect the effect of contrast to be long-lasting. Again, this is in accord with the data.

The data presented in this paper suggest that contingent duration aftereffects are different in kind from the McCollough effect. The contrast hypothesis is presented as a plausible account for the effect of adaptation. The hypothesis is compatible with existing models in the time perception literature. A more formal presentation of the hypothesis needs to be developed so that a quantitative evaluation can be undertaken.

SUMMARY

It has been shown that judged duration of tones depends on pitch (the *kappa* effect), on order of presentation (the time-order error), and on repetition (a negative duration aftereffect). Recently, a duration aftereffect contingent on pitch and a duration

aftereffect contingent on order of presentation have been described. Our results suggest that these contingent duration aftereffects differ from the McCollough effect, a color aftereffect contingent on orientation, in two ways. They have a relatively short life and they are not symmetrical, in that while a pitch-contingent duration aftereffect could be established, a duration-contingent pitch aftereffect could not. In contrast, the McCollough effect persists for days and both an orientation-contingent color aftereffect and a color-contingent orientation aftereffect have been reported. A decision theory account for contingent duration aftereffects is outlined.

REFERENCES

1. HUPPERT, F. & G. SINGER. 1967. An aftereffect in judgment of auditory duration. *Percept. Psychophys.* **2**: 544-546.
2. WALKER, J. T. & A. L. IRION. 1979. Two new contingent aftereffects: Perceived auditory duration contingent on pitch and on temporal order. *Percept. Psychophys.* **26**: 241-244.
3. WALKER, J. T., A. L. IRION & D. G. GORDON. 1981. Simple and contingent aftereffects of perceived duration in vision and audition. *Percept. Psychophys.* **29**: 475-486.
4. MCCOLLOUGH, C. 1965. Color adaptation of edge-detectors in the human visual system. *Science* **149**: 1115-1116.
5. SIROMEYER, C. F. 1978. Form-color aftereffects in human vision. *In Handbook of Sensory Physiology: (Vol. 8): Perception*. R. Held, H. Leibowitz & H. L. Teuber, Eds. Springer-Verlag, Heidelberg, West Germany.
6. MURCH, G. M. 1976. Classical conditioning of the McCollough effect: Temporal parameters. *Vision Res.* **16**: 615-619.
7. SIEGEL, S. & L. G. ALLAN. Associative bases of orientation-contingent colour aftereffects. Submitted for publication.
8. ALLEN, L. G. 1979. The perception of time. *Percept. Psychophys.* **26**: 340-354.
9. MAIEROVICH, A. M. 1983. A Duration Aftereffect Contingent on Temporal Order. Unpublished B.Sc. thesis, McMaster University, Hamilton, Ontario.
10. JAMISON, D. G. & W. M. PETRUSIC. 1975. Presentation order effects in duration discrimination. *Percept. Psychophys.* **17**: 197-202.
11. HELLSTRÖM, A. 1977. Time errors are perceptual. *Psychol. Res.* **39**: 345-388.
12. HELLSTRÖM, A. 1977. On the nature of the time-error. Report from the Department of Psychology, University of Stockholm, Supplement 38.
13. COHEN, J., C. E. M. HANSEL & J. C. SYLVESTER. 1955. Interdependence of judgments of space, time and movement. *Acta Psychol.* **11**: 360-372.
14. YOBICK, D. A. & G. SALVENDY. 1970. Influence of frequency on the estimation of time for auditory, visual, and tactile modalities: The kappa effects. *J. Exp. Psychol.* **86**: 157-164.
15. CHRISTENSEN, J. P. & Y. L. HUANG. 1979. The auditory tau effect and memory for pitch. *Percept. Psychophys.* **26**: 489-494.
16. WOODROW, H. 1935. The effect of practice upon time-order errors in the comparison of temporal intervals. *Psychol. Rev.* **42**: 127-152.
17. JAMISON, D. G. & W. M. PETRUSIC. 1976. On a bias induced by the provision of feedback in psychophysical experiments. *Acta Psychol.* **40**: 199-206.
18. JAMISON, D. G. & W. M. PETRUSIC. 1978. Feedback versus an illusion in time. *Perception* **7**: 91-96.
19. ALLAN, L. G. 1977. The time-order error in judgments of duration. *Can. J. Psychol.* **31**: 24-31.
20. JAMISON, D. G. 1977. Two presentation order effects. *Can. J. Psychol.* **31**: 184-194.

Contingent Aftereffects and Situationally Coded Criteria: Discussion Paper

MICHEL TREISMAN

*Department of Experimental Psychology
University of Oxford
Oxford OX1 3UD, England*

It is a pleasure to comment on this series of interesting and instructive papers on time perception.

Richard Schweickert's¹ paper on the application of critical path analysis to information-processing models takes the major advance in interpreting such experiments, the additive factor method first put forward by Saul Sternberg,² a step forward. Dr. Schweickert's techniques may make it possible to unravel not only simple sequential models, but also those that include concurrent parallel processes.

Drs. Wasserman, DeLong, and Larew³ examined the problem of order perception in pigeons. The study of perception in animals, as in infants, has started to take off in recent years, both because of an increased willingness to suppose that complex abilities may exist, and because of improvements in techniques for demonstrating them. This paper clearly demonstrates that pigeons make and retain discriminations of order.

Turning to the perception of temporal order in humans, George Sperling⁴ has given us a fine example of his ability for lucid analysis and incisive experimental investigation. His results are of great interest and deserve consideration at leisure.

Finally, both Donald Jamieson⁵ and Lorraine Allan⁶ have presented us with interesting observations on time-order effects. Dr. Allan has made a careful investigation of contingent aftereffects in duration judgments, giving us a great deal of information. She concludes by rejecting a parallel with the color-contingent aftereffect and suggests that the explanation for the effects she found may lie in contrast mediated by decision processes. I believe that Dr. Allan is on the right track and I would like to take this argument a little further.

"Contrast," like "assimilation," is a familiar term and is often taken for an explanatory term. But is it? I suggest that it is no more than a label for effects lying in a certain direction, and requires explanation itself. Terms like "kappa" and "tau," unfortunately, are not even descriptive.

A theory of criterion-setting on which I have recently been working⁷ offers explanations for sequential effects, of which contrast is an example, and it may be of interest to see how well such a theory can account for the present results. The theory derives sequential effects from the operation of mechanisms responsible for establishing the value of the decision criterion from trial to trial.

The theory of criterion-setting has been applied to Thurstonian models of detection, identification, absolute judgment, and magnitude estimation. These models assume that repeated presentation of a stimulus produces a normal distribution of sensory effects on a central scale, and that judgments are made by the use of decision criteria.

Let me briefly summarize the main assumptions of the theory. It addresses the

question "Why does a response criterion have one value rather than another?" and it offers the answer that there are three main mechanisms (excluding consideration of feedback) that determine the momentary values of criteria. The first, a global mechanism, establishes a long-term reference value for each criterion. This is the process implicitly assumed by signal-detection theory. It uses past experience and knowledge of global parameters, such as the overall probability of the signal and the payoff values, to determine the best reference criterion.

However, the reference criterion is not maintained without change from trial to trial, save for that produced by random noise. The reason for this is that each trial contributes short-lived information to the observer which he must apply there and then if he is to perform optimally. Consider an observer attempting to detect the presence of a stimulus. On trial i he concludes that the stimulus is present. In daily experience, objects that are truly present tend to persist and show the same or similar sensory characteristics, at least for short periods of time. Thus, the observer's criterion on the next few trials should take account of this information, which is equivalent to a short-lived increase in the probability of the stimulus. Accordingly, the observer should lower his criterion for a limited period, which will make repetition of the positive response more likely. (This corresponds to "assimilation.") By lowering his criterion in this case he "tracks" the current state of the world. To meet this requirement, criterion-setting theory proposes that each response sets up a short-lived memory trace or "indicator trace" which indicates a direction and magnitude of shift required for the criterion, and that this trace decays with time. The working criterion at any time is given by the sum of the reference criterion plus the currently nonzero indicator traces.

The tracking procedure may be put in force when the subject makes a decision, covert or overt, categorizing the input. But whether or not he does so, the sensory inputs themselves constitute a third source of information which may be used to stabilize the criterion in relation to the sensory inflow. They indicate whether it is properly centered in relation to the flux of incoming sensory inputs, or is away to one side, in which case little information will be conveyed by the responses. Thus, if the sensory input on trial i is greater than the criterion, this indicates that the criterion should shift to the right on the next trial; if the sensory input is less, the criterion should shift to the left. The extent of the shift should increase with the disparity. The overall effect of such shifts will be to center the criterion. It follows that if the sensory input is below the detection criterion on one trial, inducing a shift to the left, the probability of a detection response will be increased on the next. This constitutes "contrast." FIGURE 1 shows a reanalysis of data obtained by Tanner *et al.* in which the subject had to identify a tone as LOT D or SOFT. These data illustrate both tracking and stabilization in operation in the same experiment: if the probability of a LOT D response, $P(\text{LOT D})$, is plotted against previous responses, it is greater after a preceding LOT D response than after a SOFT one. If we plot it against the preceding stimulus magnitude, it is greater following a *soft* stimulus than following a *loud* one.

Tracking is assumed to have relatively short-lived effects, since the relevance of previous judgments may soon decline, but stabilization may produce longer-lived traces, centering the criterion in relation to a larger sample of the sensory flux.

This theory can be applied to provide explanations for the *kappa* effect, the *tau* effect, and at least some contingent aftereffects. For this purpose we add the following two assumptions: First, criteria may be situationally coded. That is, a criterion is defined not only for a given judgment, but also for particular conditions under which this judgment may be made. Then contributions to the reference criterion, and indicator traces, produced under a given condition are labeled as applying to the criterion only when employed under the same condition. Thus, the decision criterion

may have different values for different sets of circumstances. For example, if a discriminative stimulus indicates that the probability of the signal is high, this should cause the reference criterion to be lowered when that stimulus is present. This is a natural extension of the assumption that the global criterion-setting mechanism takes account of the prior probability of the signal.

The second assumption is that if a decision takes place in stages, then the evidence provided by an earlier stage may be preserved as a covert quantitative value to be employed at a later stage. But the magnitude of this quantitative value is determined

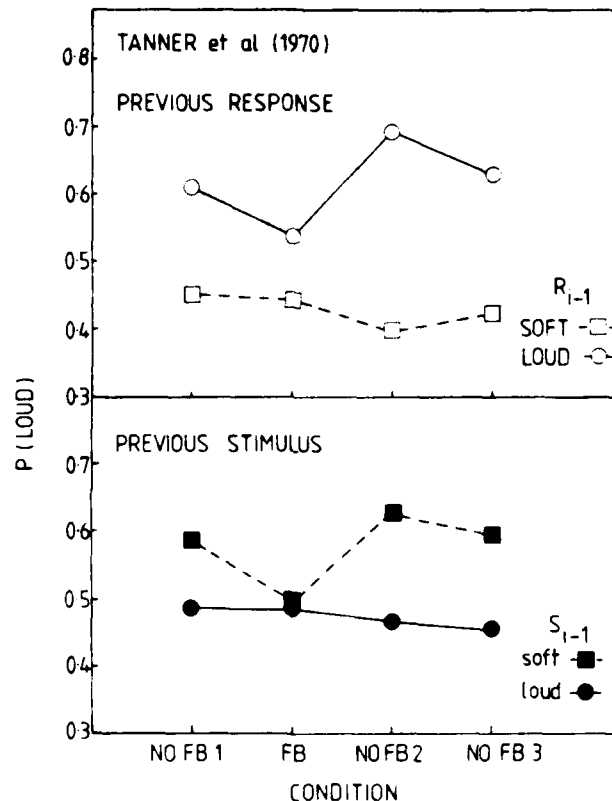


FIGURE 1. Data from Tanner *et al.*⁸ in which subjects identified a tone as LOUD or SOFT. In the upper panel the probability of the response LOUD is shown as a function of the response on the preceding trial (R_{i-1}); in the lower panel it is shown as a function of the preceding stimulus (S_{i-1}). The upper panel illustrates a positive sequential effect ("assimilation") of the preceding response; the lower panel shows a negative dependency ("contrast") on the preceding stimulus.

by a criterion whose function is to set the origin of the scale on which this value is measured. This assumption will be illustrated below.

We now examine how these assumptions may be used to build models for contingent effects.

THE KAPPA EFFECT

Cohen *et al.*⁹ exposed subjects to three lights set in order along a line and flashing in a repeated cycle. Subjects adjusted the time of the central flash to make t_c , the interval

between lights L_1 and L_2 (separated by distance d_1) equal to the interval t_2 between the flashes L_2 and L_3 (separated by distance d_2). For equal spatial separations approximately equal time intervals were produced. But if d_1 was greater than d_2 , then t_1 was less than t_2 . For example, if $d_1/d_2 = 3/1$, the mean settings gave $t_1 = 0.68$ and $t_2 = 0.82$ sec for a total cycle of 1.5 sec.

The criterion-setting model for this experiment is shown in FIGURE 2. When engaged in adjusting the time of L_2 , subjects make covert judgments of t_1 and t_2 , deciding whether each of them is too short or too long. For each judgment they use a

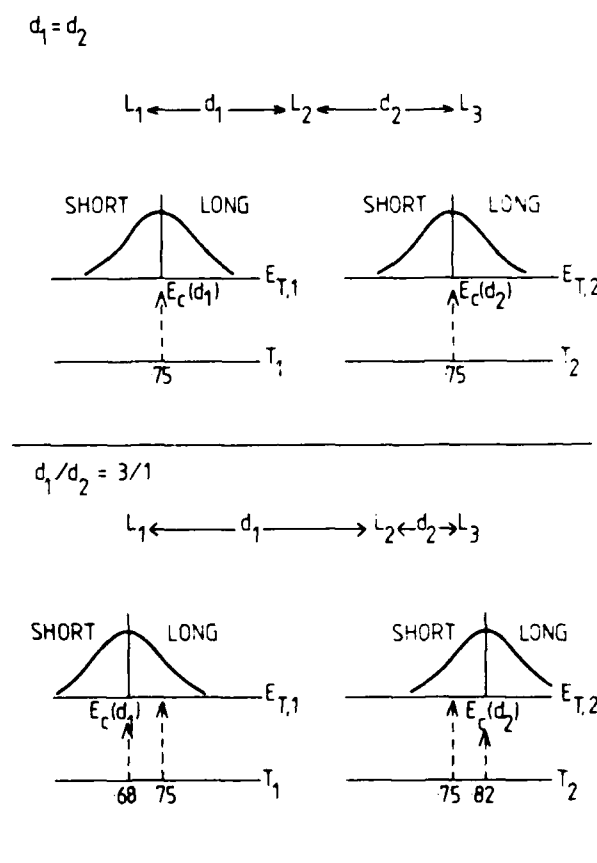


FIGURE 2. The model for the *kappa* effect⁹ given by criterion-setting theory. The *upper panel* shows the cycle of three flashing lights, L_1 , L_2 and L_3 , separated by equal distances d_1 and d_2 . The subject adjusts L_2 to flash at half the total duration. The judgments for t_1 , represented on the time dimension T_1 , and for t_2 , represented on T_2 , are shown separately. Each duration produces a distribution of central effects on the corresponding central decision axis, E_T , shown for convenience as two axes, $E_{T,1}$ and $E_{T,2}$. After each adjustment the subject decides whether the first interval is too long or too short using a criterion coded for the first distance, $E_c(d_1)$. He makes a similar covert judgment about second interval. Adjustments continue until a setting is found such that the sensory effects of the first interval are approximately equally distributed about $E_c(d_1)$, and those of the second interval about $E_c(d_2)$. Since $d_1 = d_2$, the criteria should be similar and the final judgment unbiased.

The *lower panel* applies when d_1 is three-quarters of the whole distance. The criterion associated with the long distance, d_1 , is shifted to the left, and the criterion associated with d_2 is shifted to the right. Then 0.75 seconds will seem too long for the first interval, and too short for the second. Adjustment will find a shorter interval to match $E_c(d_1)$ and a longer interval for $E_c(d_2)$.

situationally coded criterion, that is, a criterion coded by association with the first distance $E_c(d_1)$ or by association with the second $E_c(d_2)$. When $d_1 = d_2$, the criteria should be the same. But when the distances are unequal, the global mechanism will take note of this in setting the reference positions for the two criteria. In the real world movements are usually smooth: if an object moves successively through distances d_1 and d_2 , our experience is almost always that the time taken is monotonically related to the distances, especially when the movement is horizontal. Thus, for a long distance, the global mechanism should expect that the probability of a long interval will be increased and should set the criterion for detecting long intervals low; the reverse will apply when the distance is short. This is analogous to the effect of signal probability on a signal-detection criterion.

The lower panel applies the model when d_1 is long and d_2 short. The first criterion shifts to the left and the second to the right. Since the method of adjustment finds the stimulus whose mean central effect best matches the criterion, it will select t_1 less than 0.75 sec, and t_2 greater. Remember that the stabilization process tends to shift each criterion in the direction of recent sensory inputs; thus, if both start too low (giving covert LONG responses for both intervals) or too high, this will automatically be corrected.

THE TAU EFFECT

We now consider the auditory *tau* effect described by Cohen *et al.*¹⁰ These investigators presented three brief tones in succession, with an interval t_1 between the first two tones and t_2 between the second and third, and required the subject to adjust the middle tone to be intermediate in pitch. With an initial tone of 1000 Hz and a final tone of 3000 Hz, given in a 1.5-sec cycle, the middle tone was set to 2068 Hz for $t_1 = 0.5$ sec, and to 1676 Hz for $t_1 = 1.0$ sec. When the initial and final anchor tones were 3000 Hz and 1000 Hz, respectively, the middle tone was set to 1725 Hz for $t_1 = 0.5$ sec, and to 1921 Hz for $t_1 = 1.0$ sec.

This effect is formally similar to the *kappa* effect and could be explained by a similar model, although this would rest on the weaker assumption that we expect successive tones to change frequency at a regular rate. But a simpler explanation is suggested if we note that in this experiment time may function in two ways: as a processed sensory input, as the authors assume, or simply as the familiar independent variable present in all experiments. The latter is all that is required for a simpler model in which the results follow directly from the features of the tracking criterion-setting process, without requiring criteria coded for the two time intervals. This model is illustrated in Figure 3.

The figure shows the central scale for frequency, E_f , on which presentations of the anchor tones, 1000 and 3000 Hz, produce sensory distributions. It also shows the criterion E_c which would be used in identifying them. When an anchor tone is presented, the subject needs to identify it in order to perform the task; this categorization sets up a tracking adjustment to the criterion. Thus, in the figure, the categorization of each anchor stimulus sets up a corresponding indicator trace. At any moment the criterion is given by the sum of its initial value and the existing indicator traces, as the latter decay, it returns to its initial value. The method of adjustment hunts this criterion. It follows that when this method is used with $t_1 = 0.5$ sec, it will find the bisecting stimulus value further removed from the immediately preceding anchor stimulus than when $t_1 = 1.0$ sec. This is sufficient to account for the auditory *tau* effect.

CONTINGENT AFTEREFFECTS: DURATION CONTINGENT ON ORDER

Contingent aftereffects are found when judgments are preceded by an adaptation period during which the subject is passively exposed to a series of stimulus presentations which require no perceptual decisions on his part. In the absence of perceptual categorization, tracking may not occur or, since it is short-lived, its effects will be outweighed by the accumulation of longer-lived stabilization indicator traces. The passive registration of sensory inputs lends itself to the manifestation of stabilization effects. These provide the basis for a model of the contingent aftereffects described by Dr. Allan.

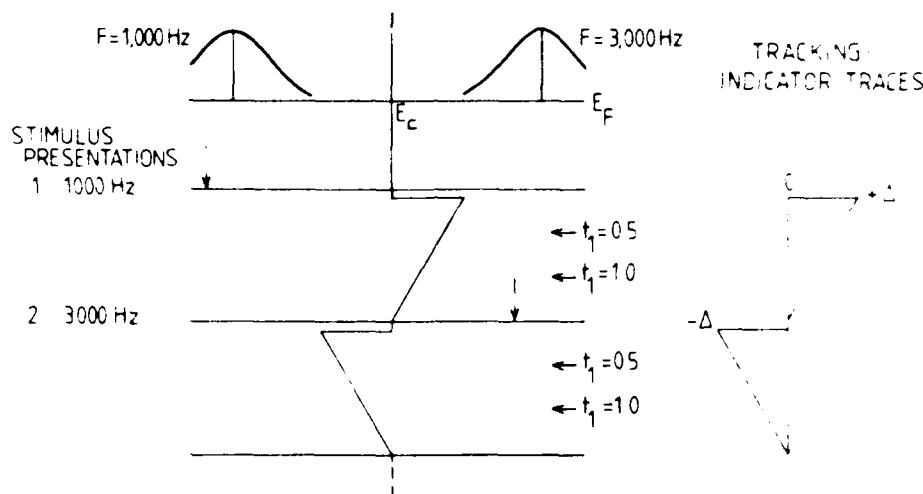


FIGURE 3. The auditory τ effect: an expression of the tracking mechanism. Repeated presentations of two anchor frequencies, 1000 and 3000 Hz, produce central effects distributed on a central scale, E_F , on which there is also an intermediate criterion, E_C . We see the effect of a presentation of the anchor stimulus, 1000 Hz. (Time proceeds downward in the figure, and the decision axis, E_F , is redrawn for each stimulus presentation.) The sensory input falls below E_C and is categorized as "1000." The tracking mechanism produces a corresponding adjustment in criterion: an indicator trace of value $+ \Delta$ is set up, and the value of this trace is added to the initial value of the criterion to give its current value. As the indicator trace decays, the current value of the criterion correspondingly returns to its initial value. In this illustration the indicator trace disappears completely before the next stimulus presentation. A second anchor stimulus is then presented, is classified as "3000" and sets up an indicator trace of magnitude $- \Delta$. This is again added to the criterion. For interpresentation intervals of 1.5 seconds, the position of the criterion when $t_1 = 0.5$, or when $t_1 = 1.0$ seconds, is indicated in each case.

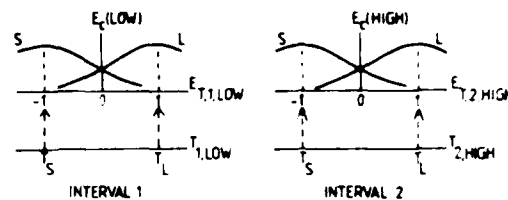
The upper panel of FIGURE 4 illustrates the preadaptation stage. The subject must decide which of two successive tones is the longer. This may be done in two stages. In the first, the stimuli are assessed as they occur, their strength being represented by quantitative outputs taken from the criterion as origin, separate coded criteria being used for interval 1 and interval 2. These outputs are compared in the second stage and determine the response: if the difference between the second measure and the first is positive, the second interval is chosen as longer, otherwise the first is selected. The upper panel assumes, for simplicity, that there is no initial bias. Then the criteria coded for the first and second intervals are the same, so that the mean sensory magnitude assigned to a given stimulus will be the same in either interval. The subject's

forced-choice decisions are determined by the difference distributions shown on the difference decision axis. (A roving standard design was used but we may disregard this since a difference distribution is calculated.) Dr. Allan defines the time-order error (TOE) as the proportion of the *LS* distribution to the left of the origin, reduced by the proportion of the *SL* distribution to the right of the origin. In the figure this is zero. In fact, for her results we should assume that the first criterion is displaced to the left or the second to the right, to account for the initial small positive error.

The effects of adaptation to short-long pairs are shown in the lower panel. During adaptation $E_c(o_1)$, the criterion associated with the first interval, is subject to the stabilization effect of the short intervals repeatedly presented in position 1. Their sensory inputs will lie well to the left of $E_c(o_1)$ and their stabilizing indicator traces will produce a shift of this criterion to the left. Similarly, the long intervals given in position 2 during adaptation will shift the criterion for this interval to the right. The sensory

BEFORE ADAPTATION

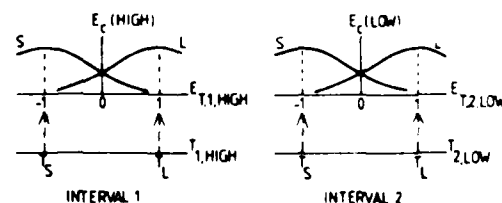
LOW-HIGH SEQUENCES



LM



HIGH-LOW SEQUENCES



HL

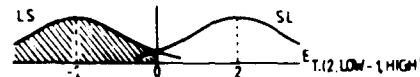


FIGURE 5A. Duration aftereffect contingent on pitch: judgments made prior to adaptation, assuming that there is no initial bias. Whether the first tone is low and the second high, or the first high and the second low, the two coded criteria, $E_c(LOW)$ and $E_c(HIGH)$, are the same in each case. Thus, the mean quantitative measure assigned to the sensory effects of a given duration, T_L or T_H , is the same whether it is the first or the second interval, and presented as a low or a high tone. Consequently, the difference distributions for the *SL* and *LS* orders will be the same and symmetrical for each tone order. $Kappa$ is given by the proportion of the *SL* distribution to the right of the origin, for the low-high tone order, plus the proportion of the *LS* distribution to the left of the origin for the high-low tone order (both shown stippled) reduced by the proportion of *LS* to the left of the origin in the first case and the proportion of *SL* to the right in the second. In this case, it is zero.

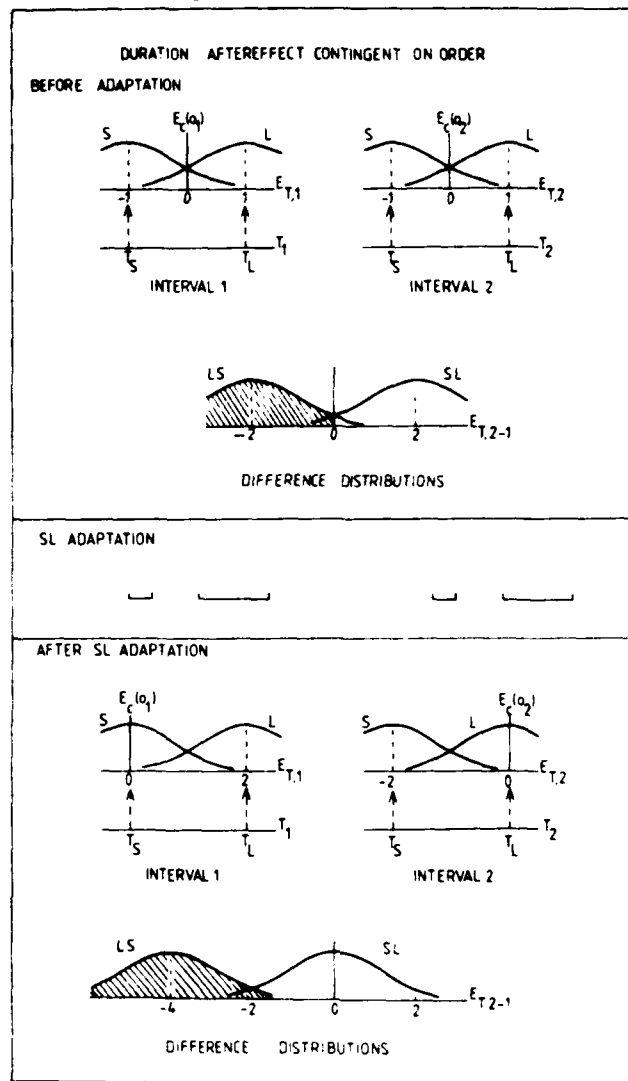


FIGURE 4. Order-contingent duration aftereffect: the model given by criterion-setting theory. Pairs of durations are presented and the subject makes a forced-choice response. The time dimension and the central effects of presented intervals are represented separately for the first and second intervals. Assuming there is no preexisting bias, the two coded criteria, $E_c(o_1)$ and $E_c(o_2)$, are identically placed. Each determines the origin on the scale of central effects, E_T , for intervals in the corresponding position. Since the origins are initially the same, the short interval (T_S) and the long interval (T_L) produce distributions of central effects with means -1 and 1 , respectively, whether these intervals occur first or second. A quantitative first-stage output is found for each interval, given by the measure of the sensory effect registered on the E_T scale for that interval. The difference between the two measures on a given trial, $E_{T,2} - E_{T,1}$, may be represented on a difference axis, $E_{T,2-1}$. The difference distributions resulting from LONG-SHORT (LS) presentations and from SHORT-LONG (SL) presentations are shown on this axis. For SL pairs the mean of the corresponding distribution is 2 , and for LS pairs the mean is -2 . The probability of a correct response to LS pairs, $P(R_{LS}|LS)$ is the area of the LS distribution to the left of the origin (stippled area) and the corresponding probability for SL pairs is the area of the SL distribution to the right of the origin. Thus, the time-order error as defined by Dr. Allan, $TOE = P(R_{LS}|LS) - P(R_{SL}|SL)$, is zero.

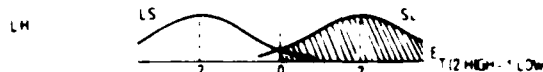
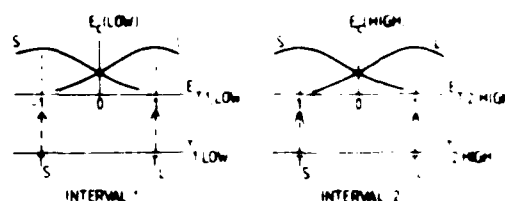
The effects of SL adaptation, to repeated SHORT-LONG pairs, on later judgments is illustrated in the lower panel. The criterion for the first interval is shifted to the left, taking the origin with it, and $E_c(o_2)$ is shifted to the right. The sensory inputs produced by the two stimuli are unchanged. But the mean sensory magnitude assigned to a presentation of T_L in the first interval is now 2 , not 1 , because of the criterion shift to the left, and in the second interval it is now 0 , not 1 , because the origin has moved to the right. Thus, the means of the difference distributions are changed for each order of presentation. Both are shifted to the left and the TOE is increased.

forced-choice decisions are determined by the difference distributions shown on the difference decision axis. (A roving standard design was used but we may disregard this since a difference distribution is calculated.) Dr. Allen defines the time-order error (TOE) as the proportion of the LS distribution to the right of the origin, reduced by the proportion of the SL distribution to the right of the origin. In the figure this is zero. In fact, for her results we should assume that the first interval is displaced to the left or the second to the right, to account for the initial small error.

The effects of adaptation to short-long pairs are shown in the lower panel. During adaptation $E_c(o_1)$, the criterion associated with the first interval, is subject to the stabilization effect of the short intervals repeated in position 1. Their sensory inputs will lie well to the left of $E_c(o_1)$ and their stabilizing indicator traces will produce a shift of this criterion to the left. Similarly, the long intervals given in position 2 during adaptation will shift the criterion for this interval to the right. The sensory

BEFORE ADAPTATION

LOW-HIGH SEQUENCES



HIGH-LOW SEQUENCES

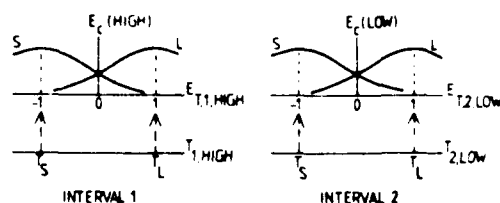


FIGURE 5A. Duration aftereffect contingent on pitch: judgments made prior to adaptation, assuming that there is no initial bias. Whether the first tone is low and the second high, or the first high and the second low, the two coded criteria, $E_c(LOW)$ and $E_c(HIGH)$, are the same in each case. Thus, the mean quantitative measure assigned to the sensory effects of a given duration, T_1 or T_2 , is the same whether it is the first or the second interval, and presented as a low or a high tone. Consequently, the difference distributions for the SL and LS orders will be the same and symmetrical for each tone order. $Kappa$ is given by the proportion of the SL distribution to the right of the origin, for the low-high tone order, plus the proportion of the LS distribution to the left of the origin for the high-low tone order (both shown stippled) reduced by the proportion of LS to the left of the origin in the first case and the proportion of SL to the right in the second. In this case, it is zero.

inputs given by T_S and T_L are unchanged. But the quantitative measures they give rise to in each interval are shifted because of the shifts in origin (the order-coded criteria) on the E_T scale for each interval, and in consequence the LS and SL difference distributions are each moved to the left. This will manifest as an increase in TOE . In the same way, LS adaptation would produce a decrease in TOE .

DURATION AFTEREFFECT CONTINGENT ON PITCH

In her fourth and fifth experiments Dr. Allan adapted subjects to short and long intervals (200 and 600 msec) given as low and high tones (600 and 900 Hz). The

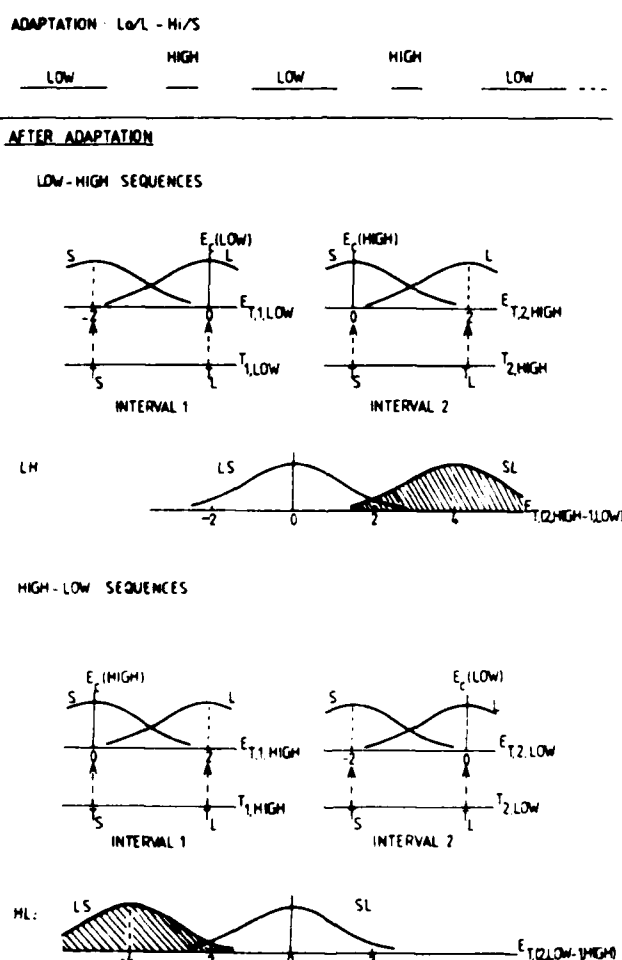


FIGURE 5B. Duration aftereffect contingent on pitch: Lo/L-Hi/S adaptation and the state of the criteria after such adaptation. The criterion coded by association with the low tone, $E_L(LOW)$, has been exposed to long intervals during adaptation. It has shifted to the right, so that in this illustration the origin now coincides with the mean of the T_L distribution. This holds for either tone sequence during judgment. Similarly, $E_L(HIGH)$ is shifted to the left in both cases. Although the sensory inputs produced by T_L and T_S have not changed, they will now both be assigned lower values when presented as low tones and higher values when presented as high tones. Consequently both difference distributions are shifted to the right for the low-high tone sequences and to the left for the high-low sequences. Consequently, $kappa$ is increased.

subject was subsequently required to decide which of two tones (one at 600, the other at 900 Hz) was the longer. An effect was found which Dr. Allan defines as

$$\begin{aligned} \text{kappa} = & [P(R_{SL} | SL \text{ and } LH) + P(R_{LS} | LS \text{ and } HL)] \\ & - [P(R_{SL} | SL \text{ and } HL) + P(R_{LS} | LS \text{ and } LH)]. \end{aligned}$$

Adaptation to low-frequency long tones and high-frequency short tones increased *kappa*; the reverse adaptation reduced it. These results can also be explained by stabilization effects on coded criteria, the criteria in this case being coded by association with the low or high frequencies. The model is illustrated in FIGURES 5A and B.

Prior to adaptation, and assuming no initial bias, the criteria coded by pitch are the same, giving symmetrical difference distributions and a zero value for *kappa*. If low-frequency long-duration intervals are given during adaptation, these will act only on the criterion coded for association with low frequency. Since they are long, they will cause stabilization shifts of $E_c(\text{LOW})$ to the right. Equivalent effects will be produced by other adaptation combinations. The figure demonstrates how *Lo/L-Hi/S* adaptation shifts the criteria selectively and consequently causes the difference distributions to shift to the right for low-frequency-high-frequency sequences and to the left for high-low sequences. This gives a positive value of *kappa*. Similarly, it can be shown that *Lo/S-Hi/L* adaptation will cause a reduction in *kappa*.

Prior to adaptation, Dr. Allan's subjects showed a positive value of *kappa*. This may possibly reflect a bias in the environmental noises we experience, low-pitched rumbles tending to be longer than higher-pitched squeaks.

It appears that criterion-setting theory, extended by the hypothesis of situational coding of criteria, can account for the major intersensory effects involving time. This result has a number of interesting implications.

1. If criterion-setting theory can account for the present contingent aftereffects, it is possible that it may also account for many of the contingent aftereffects found in other modalities, such as the color-contingent aftereffect.¹¹ This would relate such phenomena directly to sequential dependencies in general and so provide a more parsimonious account than do theories that require the assumption that neural feature detectors are the site of adaptation or theories that postulate special types of learning.

2. Whereas the color-contingent aftereffect may endure days or weeks,¹² Dr. Allan found that her order-contingent aftereffect did not last 25 minutes. The explanation for this difference may lie in situational coding: if a coded criterion is modified only in the situation it is intended for, then its life may be a function of the frequency with which it encounters that specific situation and is modified in it. The natural environment frequently presents us with sequences of sounds of varying durations, which may expose criteria coded for order to noisy influences, whereas the bar and color patterns used in inducing the color-contingent aftereffects are rarely seen under natural conditions.

3. In her final two experiments, Dr. Allan failed to find an effect on pitch contingent on the duration of adapting stimuli. Thus, this experiment provides no evidence that pitch criteria may be coded by temporal duration. We have also seen above that the auditory *tau* effect of Cohen *et al.*¹⁰ does not require coded criteria for its explanation. These observations raise the possibility that duration cannot readily be used to code pitch criteria and perhaps criteria on other modalities. The presence or absence of contingent aftereffects may provide us with evidence on relations between modalities which allow inputs in one to affect sensory decisions in another.

4. The present analysis provides a basis for distinguishing between different types

of cross-sensory effects on decision. Thus we can contrast Cohen *et al.*'s κ , which we have attributed to the global criterion-setting process, and the finding described by Dr. Allan, which may be mediated by stabilization of coded criteria.

REFERENCES

1. SCHWEICKERT, R. J. 1984. The representation of mental activities in critical path networks. This volume.
2. STERNBERG, S. 1969. The discovery of processing stages: Extensions of Donders' method. In *Attention and Performance: II*. W. G. Koster, Ed. North-Holland, Amsterdam.
3. WASSERMAN, E. A., R. E. DE-LONG & M. B. LAREW. 1984. Temporal order and duration: Their discrimination and retention by pigeons. This volume.
4. SPERLING, G. Temporal order of brief visual events. Presented at this conference, but not submitted for publication.
5. JAMIESON, D. G., E. SIWINSKA, M. F. CHEESMAN & B. E. VARAS. 1984. Timing perturbations with complex auditory stimuli. This volume.
6. ALLAN, L. G. 1984. Contingent aftereffects in duration judgments. This volume.
7. TREISMAN, M. 1984. A theory of criterion setting with an application to sequential effects. *Psychol. Rev.* **91**: 68-111.
8. TANNER, T. A., J. A. RAUK & R. C. ATKINSON. 1967. Signal recognition as influenced by information feedback. *J. Math. Psychol.* **7**: 259-274.
9. COHEN, J., C. E. M. HANSEL & J. D. SYLVESTER. 1955. Interdependence in judgments of space, time and movement. *Acta Psychol.* **11**: 360-372.
10. COHEN, J., C. E. M. HANSEL & J. D. SYLVESTER. 1954. Interdependence of temporal and auditory judgments. *Nature (London)* **174**: 642-644.
11. MCCOLLUGH, C. 1965. Color adaptation of edge-detectors in the human visual system. *Science* **149**: 1115-1116.
12. STROMMEYER, C. F. 1978. Form-color aftereffects in human vision. In *Handbook of Sensory Physiology*, Vol. 8. R. Held, H. Leibowitz & H. L. Teuber, Eds. Springer-Verlag, Heidelberg.

Introduction

ALAN M. WING

*Medical Research Council
Applied Psychology Unit
Cambridge CB2 2EF, England*

Of the following four papers, three—those of Stelmach and Requin and Semjen, and my own—are concerned with the temporal patterns produced in sequences of movements, while the other, that of Jones, describes a pitch-recognition task in which the stimuli are presented in different temporal patterns. All four papers use human subjects, while my own, with Keele and Margolin, takes advantage of deficits in performance resulting from a neurologic disorder to explore a theoretical model developed on the basis of the more usual “normal” population of college students.

The papers by Jones and Requin and Semjen return to a theme taken up in other research reported at this conference: Physically distinguishable dimensions of the stimuli (the interstimulus interval of Schweickert) or of the movements (the interresponse interval of Eisler) interact in their effects on performance. The internal representation of a series of movements and the consequences of timing of the movements is addressed by three of the papers: in the paper by Stelmach the timing is incidental to the production of handwriting, but in the work of Requin and Semjen and Wing *et al.* the subjects have to produce well-regulated interresponse intervals. All three papers examine the intervals between excessive movements, either in terms of means (Stelmach and Requin and Semjen) or in terms of variances (Wing *et al.*). In addition, the papers by Stelmach and Requin and Semjen attempt to characterize the nature of separation for movements using the latency with which the sequences are initiated.

PREVIOUS PAGE
IS BLANK

Motor Programming and Temporal Patterns in Handwriting^a

GEORGE E. STELMACH,^b PATRICIA A. MULLINS,^c AND
HANS-LEO TEULINGS^d

^b*Motor Behavior Laboratory
University of Wisconsin
Madison, Wisconsin 53706*

^c*Committee on Cognition and Communication
University of Chicago
Chicago, Illinois 60637*

^d*Psychological Laboratory
University of Nijmegen
Nijmegen, the Netherlands*

Over the past decade there has been considerable interest in the process by which motor commands pass from some abstract representation in the brain to a functional code at the muscular level.¹⁻⁴ The term most frequently used in conjunction with this aspect of movement specification is that of a motor program. Henry and Rogers¹ and Keele⁴ purported the motor program to be a specification of parameters of action such as force, velocity, duration, and sequencing of the involved muscles. Over the years this idea has changed so that today a motor program is thought of as an abstract non-muscle-specific representation of motor acts.

Recently, an increasing number of studies have investigated motor programs in complex motor tasks, such as the production of words,^{5,6} typewriting,⁷ and piano playing,⁸ in an attempt to describe the characteristics of these programs. Handwriting, involving the coordination of orthogonal muscle systems of the forearm, hand, and fingers, with intricate timing relationships, is another skill that is proving to be useful for studying the nature of motor programming.

As described from a motor control perspective, handwriting can be viewed as the production of flexion and extension movements of the thumb, index, middle, and ring fingers, and abduction and adduction of the hand around the wrist joint. For the most part, vertical strokes of letters are produced by thumb-finger flexions and extensions, and horizontal strokes are produced by hand abduction and adductions. A steady rotation about the shoulder produces left-to-right progression. Despite the rich theoretical motor control aspects of handwriting, most research until recently has concentrated on handwriting characteristics and pathologic abnormalities.⁹⁻¹³ However, the widely quoted description of the handwriting simulated by Hollerbach,¹⁴ the expanding research effort from the Psychological Laboratory at the University of Nijmegen,¹⁴⁻¹⁶ and the time-space constancy data by Viviani and Terzuolo¹⁷ have all increased interest in handwriting as a skill for gaining insight into the representation of complex actions and the structure and organization of motor commands.

^aThis research was conducted at the University of Nijmegen and was partially sponsored by the Graduate School Research Committee of the University of Wisconsin, by the Scientific Affairs Division of NATO, by the Senior Fulbright Scholar Program, and by the Netherlands Organization for the Advancement of Pure Research.

Recent research from Nijmegen^{14, 16, 20} has clearly shown that handwriting movements, often considered to be continuous for letters and words, can be decomposed into individual strokes. The resulting dynamic description of size, velocity, and duration parameters of the writing trace is used to help understand the control parameters of the motor output. Using such measures, Viviani and Terzuolo¹⁷ found in handwriting an invariant structure for space-time parameters, leading them to conclude that the basic unit of organization for motor output corresponds to the sequence that is specified rather than letter-to-letter transitions. The homotetic behavior in the time domain for handwriting is captured by the tangential velocity when subjects are required to intentionally modify their writing speed while keeping constant the size of the letter. When the total duration of the writing movement changes, instantaneous values of the velocity change proportionally in such a way as to leave invariant the ratios among the times of occurrence of the major features of the trajectory profile. Similar trajectory profiles are also observed when changes in writing size are required.^{17, 21} This observation led Viviani and Terzuolo to conclude that the observed homotetic behavior in the time domain is a general organizational principle of learned movements.

Along a different dimension, Raibert⁶ has shown that handwriting possesses striking similarities when writing with the left or right hands, right arm, pen taped to the right foot, or pen held in the mouth. The observation suggests that there is a common memorial representation for all of the writing behaviors that is implemented regardless of limb and muscles involved.⁵ Keele,¹ attempting to relate these data to Hollerbach's¹³ spring model, suggests that if, in each of Raibert's handwriting examples the muscles are organized into relationships that operate at right angles to each other, then the same time patterns could be applied and the output, except for size, would be very much the same.

These studies suggest that cursive handwriting may result from the relative timing of sets of muscles in a generalized force-time relationship. A program for a letter could then be thought of as a specification of the phasing between coupled oscillations in the horizontal and vertical direction with transitions from one letter to another occurring when the phases are changed by some underlying timing component. The unit that is specified would be in this sense a complete cycle at one phase setting of the agonist-antagonist muscle pairs.¹

In search of a unifying principle underpinning these various hypotheses, two fundamental issues are investigated herein, namely, the generality of the invariant characteristics of the relative timing of a handwriting sequence, and the nature and specification of a unit of the motor program for handwriting. The first experiment to be reported examines the susceptibility to modification of the widely acclaimed space-time invariance reported by Viviani and Terzuolo,¹⁷ in a prepared handwriting sequence. Stroking characteristics in situations where a given sequence is prepared and executed are compared to those where the prepared sequence must be modified at the time of an imperative response signal. Modifications of a prepared allograph required the subjects to either increase handwriting size (3 cm) or decrease it (1.5 cm). Preparation state was manipulated by expectancy information that created a bias toward preparing one of the two handwriting sizes. This type of paradigm was first used by La Berge, van Gelder and Yellott¹⁸ and has been recently employed by Larish and Stelmach¹⁹ and Stelmach and Teulings²⁰ to examine the restructuring aspects of motor programming processes.

At issue in the present experiment is whether the space-time invariance phenomenon operates differentially under advance-planning and parameter-restructuring situations. Because size is the only variant between probability conditions in the experiment, the main difference between conditions is the "reparameterizing" or readjustment of force-time parameters required for proper execution.^{20, 22} Heretofore,

the space-time invariance has only been observed in situations where the subject could fully prepare the execution of writing different sizes. One of the issues to be explored in this experiment is whether the effect of change in size prescription is limited to the initial stroke or whether the context of the size preparation will alter the duration of the stroke characteristics throughout the allograph. If learned handwriting movements are represented in some abstract code and retrieved as a unit, it may be posited that in the restructuring situation the subject will have difficulty adjusting size parameters throughout the handwriting sequence.

EXPERIMENT 1

Method

Subjects

Subjects were 13 right-handed male and female psychology students from the University of Nijmegen. They were either paid or given class credit for participation.

Apparatus

The writing movements were recorded by a computer-controlled digitizer (Vector General Data Tablet DTI). The position of the tip of the electronic pen, expressed in horizontal and vertical coordinates with a combined RMS error better than 0.2 mm, was sampled at a rate of 200 Hz.²⁰ The pen tip was an ordinary ballpoint refill and the subject wrote on a sheet of paper fixed on the digitizer surface. The digitizer was positioned such that the subject's individual writing slope was parallel to the horizontal axis of the digitizer. Direct vision of the writing hand was eliminated by the placement of a shield above the writing surface.

A display (Vector General Graphics Display Series 3 Model 2DS) was positioned at a distance of 125 cm directly in front of the subject at eye level and it allowed the tachistoscopic presentation of stimuli (the stimuli were built up within 1 msec).

Procedure

Each trial began with a buzzer and after a 1-second delay, a writing stimulus was displayed for 100 msec. One allograph was presented (*h_{ye}*) and depending on the probability condition, the subjects were required to write it in one of two sizes (large or small). The subject's task was to initiate a response to the writing stimulus as fast as possible with minimal error. All pen movements were recorded and stored and 1 second later the recorded writing trace, the stimulus, and the reaction time were displayed to the subject for 2 sec. The fed-back writing trace served also as a means to instruct the subjects in the correct writing size. As such, the traces of the letters were fitted between two horizontal lines representing the correct writing size of 1.5 or 3 cm. These lines were also useful to inform the subject that s/he was writing horizontally and correctly.

The subjects were familiarized with the writing task and given sufficient practice (100 trials of each allograph size) until they had no difficulty. Then each subject participated in a block of 55 trials where the type of allograph displayed and written

altered between *h*ye large (3 cm) or small (1.5 cm). As an aid to preparation, the highly probable allograph size was continually visible to the subject during a block of trials and when the subjects heard the buzzer (1 sec prior to "go" signal), they were to prepare the displayed handwriting response and write it as fast as possible when the allograph was flashed on the display. The distinguishing feature between blocks was the designation of which allograph size appeared most frequently. Within a block of trials, 40 trials consisted of the flashed allograph matching the displayed allograph (80%), 10 where it did not (20%) and 5 that were catch trials (no target signal). Depending on the block, the subjects participated in sessions where they either executed *h*ye large 80% of the time, but on 20% of the trials had to switch to *h*ye small, or else they executed *h*ye small 80% of the time, but had to switch to *h*ye large on 20% of the trials. Thus, for each allograph size there was a highly prepared execution (80% trials) and one that was executed only after it was restructured (20% trials).

It was stressed that all up and down movements had to be made clearly, that at the start of a trial the pen had to be in contact with paper, and that the pen should not be lifted during the writing movement. The subjects were instructed to perform the writing task as fast and as accurately as possible. It was stressed that the subject should make optimal use of the probability information by preparing the most probable allograph on every trial and switching to the alternative sequence only if the stimulus display required it. Presentation order of allographs and probability levels were balanced over subjects. The first five trials were exclusively the most probable stimulus, and no two successive catch trials with the least probable grapheme occurred.

Response Analysis

The vertical coordinate as a function of time was differentiated and filtered at 16 Hz, yielding the vertical velocity, and time marks were determined where the vertical velocity changed sign (that is, a downward movement is passed into an upward movement or vice versa). The allographs were chosen such that the movement from one time mark to the following could be regarded as one stroke.¹³ In addition to the intervals between these time marks, the velocity, the net length, and the direction of the strokes were also determined for the first eight strokes of the allographs. The time marks, as calculated by a computer algorithm, were made visible so that the analysis of each trial could be verified.¹⁶ If the algorithm did not work perfectly, the time marks could be readjusted by hand, using a moving cursor controlled by a 10-turn potentiometer.²⁰ If the response obviously contained an execution error, the trial was excluded from analysis.

Results

Reaction Time

The mean reaction times (RTs) within each handwriting size and probability level were determined and then compared, and these reaction times for the four conditions are reported in TABLE 1. Inspection of the table reveals that for both handwriting sizes the RTs are considerably faster in the high-expectancy conditions (80%) compared to those of low expectancy (20%). By employing a one-tail sign test and using subjects as uncorrelated variables, we observed that both handwriting sizes were initiated

TABLE 1. Mean Reaction Times (msec) for *hye* Written Small or Large under Two Probability Levels

	Small	Large
<i>hye</i> -20% condition	383	362
<i>hye</i> -80% condition	317	301
Mean difference	66	61

significantly faster (large, 301 versus 362 msec, $p < .05$; small, 317 versus 383 msec, $p < .01$) in the advanced-planning conditions, indicating that the subjects were indeed using the probability information to prepare the expected handwriting sequence.

It is also worthy of note that when comparisons in RTs are made between handwriting sizes, the RTs for the smaller-sized writing are consistently longer, regardless of expectancy level. For the 80% condition, the difference was 16 msec, and for the 20% condition it was 21 msec. While the exact locus of this effect is unknown, it most likely due to differences in the force-time prescription at the muscular level. Handwriting consists of a set of orthogonal muscle control systems with intricate timing relationships, and in this case, when large amounts of force are required, the wrist joint and its constituent muscle groups probably play a larger role in the initial stroke than in the smaller-execution size, where finger movement dominates. Similar findings were reported by Bonnet, Requin and Stelmach,²² who found that specification of short dorsal-flexion foot movements took longer to initiate than substantially longer ones.

Stroke Characteristics

The average length per stroke for the production of the two allograph sizes is reported in TABLE 2. It is easily seen that the execution of the two allograph sizes produces substantially different lengths per stroke and that, in general, having to switch to a different size allograph had little effect on stroke length. This indicates that regardless of the type of preparation, the subjects are able to accomplish the switch to produce the required lengths, with the differences being consistent throughout the first eight handwriting strokes. However, close inspection of the first stroke for each probability condition reveals that the type of preparation alters the produced length of

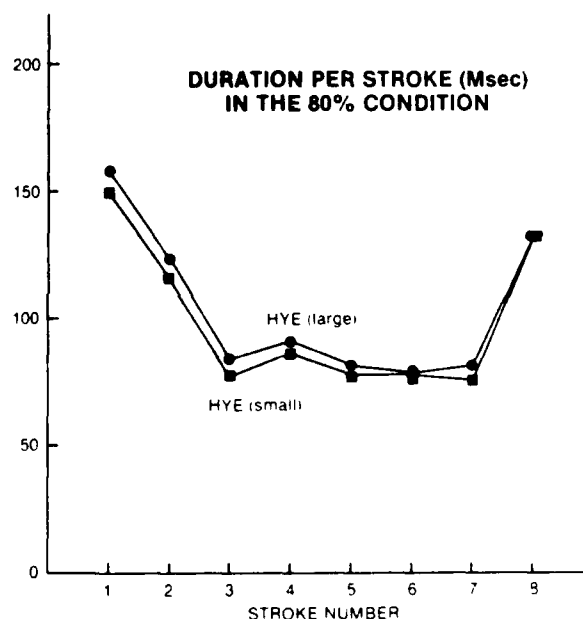
TABLE 2. Mean Lengths and Velocities per Stroke for the *hye* Allograph Written Small and Large for each Probability Level

	Stroke Number							
	1	2	3	4	5	6	7	8
Length (mm)								
<i>hye</i> -large 20%	1.47	1.53	0.61	0.42	0.56	0.31	0.55	1.21
<i>hye</i> -small 20%	.97	.81	.34	.23	.31	.16	.32	.68
<i>hye</i> -large 80%	1.62	1.58	.66	.45	.58	.31	.58	1.21
<i>hye</i> -small 80%	.84	.74	.34	.23	.31	.16	.30	.70
Velocity (cm/sec)								
<i>hye</i> -large 20%	12.16	18.71	8.14	6.88	6.55	5.57	6.12	12.97
<i>hye</i> -small 20%	9.05	10.41	4.45	3.97	3.78	6.10	3.58	8.07
<i>hye</i> -large 80%	14.52	20.21	9.17	7.76	7.28	5.83	6.61	13.92
<i>hye</i> -small 80%	7.58	9.97	4.47	4.08	3.64	3.01	3.33	8.08

the first stroke. Under conditions where the subject prepared for a small stroke, but then had to execute a large one, the length of the first stroke is shorter than when he or she prepared to write a large stroke and then executed it. A similar context effect, but in the opposite direction, is found for the first stroke when subjects prepared to write large, but had to switch to execute a small stroke. Both of these context effects are significant ($p < .05$).

Velocities per stroke are also reported in TABLE 2 and it is apparent that with an increase in length per stroke there is an associated increase in velocity, the longer stroke lengths possessing large velocities. These increases in velocity remain fairly constant across the eight strokes regardless of the type of preparation. There also seems to be a proportional increase in velocity for the longer strokes within an allograph, as is seen by making comparisons between length and velocity on a stroke-by-stroke basis. Moreover, when we examine the lengths of the first strokes, we see a context effect in the first stroke for velocity. When the subjects prepared for a large stroke and then executed it, they did it faster than when they had to switch to execute a small stroke, and vice-versa.

FIGURE 1. Duration per stroke in msec for the 80% condition plotted as a function of stroke number for *hye*-large and *hye*-small.



While the observed stroke parameters for length and velocity are along the lines suggested by a space-time invariance,¹⁷ the data on durations per stroke are rather informative. FIGURE 1 plots the durations per stroke for small versus large in the 80% condition, where the subject executed the prepared size. Statistical comparisons between the durations per stroke revealed that there are no differences between small and large sizes, supporting the well-known time-constancy invariance often observed in handwriting.¹⁷ FIGURE 2 plots the durations per stroke for the "reparameterizing" condition (20%) as a function of large and small sizes. In contrast to the results in FIGURE 1, there are substantial differences between writing size for stroke durations across the eight strokes plotted. All of these differences are significant at the $p < .01$ level and suggest that the force-time phasing is not normally specified. Yet it can be seen by the shape of the curve that the relative ratios among the times of occurrence of the individual strokes is invariant.

In comparing across figures it can be seen that most of the duration effect is due to differences in the execution of the large allograph. The duration for large strokes in the 20% condition is considerably longer, whereas there is little difference in the 80% condition. This implies that it is more difficult to increase the force required when the preparation is for a small size, than when a decrease in size is required. Remarkably, this duration effect remains for the entire eight strokes, suggesting that there is an overall force-time prescription that modulates the entire allograph.

Discussion

The foregoing data suggest that the type of preparation (80% versus 20%) within a given size affects the execution of the first stroke of an allograph, as seen in the significant differences observed in velocity and length parameters in TABLE 2. This

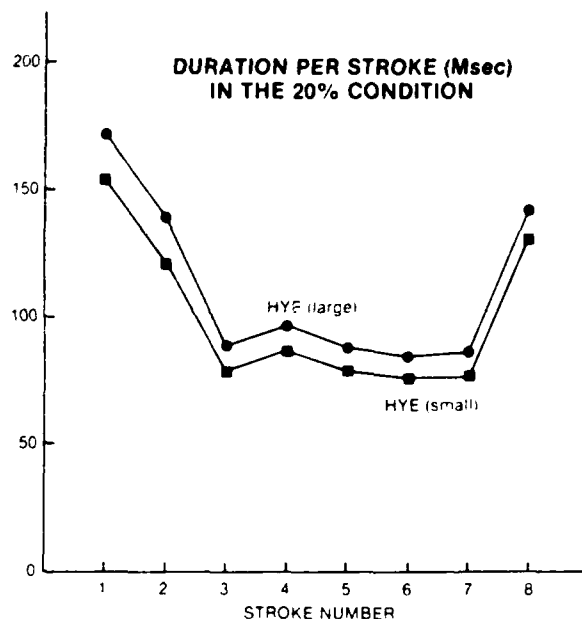


FIGURE 2. Duration per stroke in msec for the 20% condition plotted as a function of stroke number for *hve*-large and *hve*-small.

indicates that not all effects of programming are indexed by the reaction time. The approximate cost of "reparameterizing" the size prescription for both writing sizes was 60 msec. Apparently, this additional time was not sufficient to get the force-time relationship organized so that the sequence could be normally executed. An examination of advance-planning effects as seen by comparing length and velocity parameters within a handwriting size reveals that the observed preparation effects are limited to the first stroke. One interpretation of this part of the data is that the motor programming in these handwriting tasks is limited to the first stroke, with the remainder of the programming done on line. This observation is similar to that proposed by Hulstijn and van Galen,¹⁵ who found that interdigit intervals appear not to be influenced by the sequence length, thereby reflecting a local programming process for the execution of each individual digit. Our data, interpreted as such, are somewhat at odds with those of Sternberg *et al.*,⁷ who proposed that an entire sequence of words (speech) or characters (typewriting) are being programmed prior to the initiation of the first response and

that a global search process is necessary for the execution of each individual response in the sequence.

A slightly different picture of motor programming is obtained when one compares large versus small writing sizes at each of the expectancy conditions (80% versus 20%). Under the "reparameterization" condition (that is, when a movement sequence is prepared under a specific force-time relationship and this preparation has to be restructured) it was found that the original preparation interferes with the new one. The surprising aspect of this finding is that the lack of adequately changing the durations in the handwriting sequence for the 20% condition remains throughout the execution of the allograph and the timing ratios of successive elements are preserved. As for the debate mentioned earlier, over the nature of motor programming and the type of units involved, these data suggest that motor programming for handwriting may be characterized as a series of horizontal and vertical movements which are modulated by some underlying timing component involving the entire sequence. A few years ago, Wing²³ came to essentially the same conclusion from his work on durations between stroke segments and suggested that the basic program unit being timed was not a single stroke, but that an underlying metronomic process was involved.

As can be seen from the evidence presented thus far, a programming unit for handwriting has been described variously as an agonist-antagonist phase cycle,¹¹ a metronomic process,²³ word,¹⁷ and a letter.¹⁵ Further, since it has been found that increasing the number of line segments in the production of a grapheme increases the reaction time,¹⁴ there has also been some speculation that individual strokes may be represented as programming units. If we are to interpret work on motor programming in handwriting, it is clear that what constitutes an elementary unit of a program must be defined. To this end, two experiments were designed to uncover whether the most fundamental strokes, straight or curved, could function as programming units in a paradigm similar to that used by Sternberg *et al.*^{7,8} The number of strokes was manipulated in order to test whether the length of the sequence had an effect upon the RT interval. A linear increase in RT and a quadratic increase in duration would then indicate advance planning of a sequence of strokes, with the unit being the individual stroke.

The second and third experiments employed a modified simple reaction-time paradigm with a handwriting response. Figures consisting of 1-5 line segments ($\backslash \vee \eta \omega \mathbb{W}$) and 1-6 line-curve permutations ($\backslash \cup \cup \omega \omega \omega$) were presented individually in random order. The subjects' task was to reproduce the prespecified figure, completing it as quickly as possible after receipt of an imperative signal. If the unit of programming in handwriting is not the stroke, but is one complete cycle at a particular phase setting,^{1,13} with letters emerging as modulations of underlying oscillation patterns, or if it is a letter, then differential programming effects should occur not only among items in one experiment, but also between items when comparing both experiments.

EXPERIMENT 2

Methods

Subjects

Subjects were four right-handed male and female psychology students drawn from the same population as Experiment 1.

Apparatus

The apparatus and experimental configurations were basically the same as in Experiment 1, except that direct vision of the hand was possible, and that the vertical coordinates were filtered and differentiated at 32 Hz.

Procedure

A typical trial proceeded as follows: A line-segment figure (for example, π) appeared on the display screen for 500 msec and then went off. After a 2500-msec delay, which subjects were encouraged to use for response preparation, two 100-msec warning beeps (1000 Hz) occurred, 500 msec apart. They were followed 500 msec later by a third 100-msec beep having a higher frequency (2000 Hz) tone. This was the signal to write the figure, completing it as rapidly and accurately as possible. All movements of the pen were recorded and stored and 500 msec later a feedback trace of the response was displayed and the next trial began automatically. In order to avoid anticipation, on 16% of the trials the third tone did not sound, indicating a catch trial on which the subject was not to respond.

Summary information was given at the end of each block in terms of the number of errors out of the total number of real trials, the average movement time in milliseconds between the signal and the end of the writing response, and a score. The score was based on time, errors, and standard deviation, encouraging the subjects to optimize the speed-accuracy trade-off. The subjects were familiarized with the writing task and practiced 375 trials, attaining a high level of performance. They were then presented with four blocks of 75 trials, 12 repetitions of each stimulus item per block plus catch trials. Each block was pseudorandomized with a constraint of no two identical successive items. The subjects were encouraged to write the figures in a consistent size with all up and down movements being made clearly and precisely. They were informed that the pen had to be in contact with the paper at the start of a trial and could not be lifted during the writing movement.

EXPERIMENT 3

The same subjects participated in this experiment as Experiment 2, with the order of presentation counterbalanced. The stimulus items were the only change in this experiment. Six different figures were used here ($v u w u w$). After 450 practice trials, subjects were presented with four blocks of 90 trials each. This experiment was identical in all other respects to Experiment 2.

Results

Reaction Time

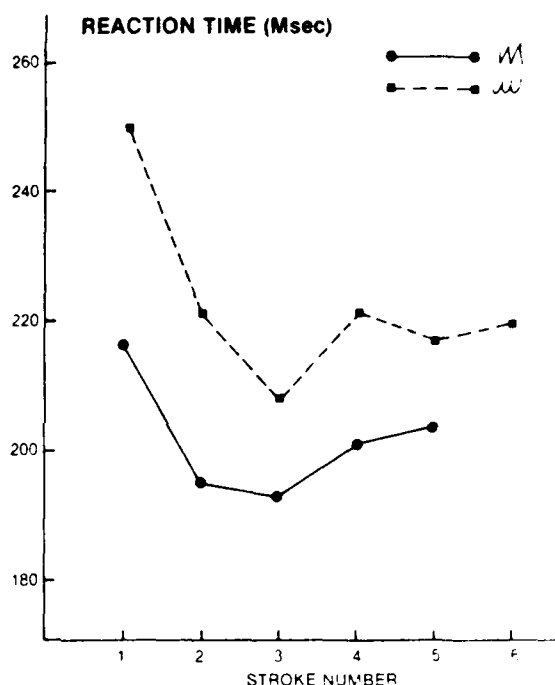
The mean reaction time for each sequence length, averaged over subjects and over sessions, is shown in FIGURE 3. Results for the stimulus items in Experiment 2 and Experiment 3 are considered both separately and in relation to each other.

Visual inspection reveals some similarity in the shape of functions, with the most striking effect being the increase in RT for the first stroke in Experiment 2 and the first

2 strokes in Experiment 3. This finding has been demonstrated in previous work on handwriting,¹⁵ where a possible explanation was put forth that for short movements the stop (which for handwriting measurements must be very precise) is programmed before the movement starts. Sternberg *et al.*⁷ discuss the element invariance requirement in which the equivalence of single, beginning, interior, and terminal elements must be assumed in order to interpret the latency function for sequences starting with $n = 1$. If this requirement is not met, and indeed it may not be in handwriting, given the bidirectional phase shifts, our attention must be restricted to performance functions for $n \geq 2$ (or $n \geq 3$, in Experiment 3, where u may be considered a single element).

The best-fitting straight line in Experiment 2 was calculated for the remaining points and a trend analysis on the linear component was performed. This analysis

FIGURE 3. Reaction time in msec for each of the line and line curve segment experiments plotted as a function of stroke number.



yielded a slope of 5.2 msec of the fitted latency function $F(1, 6) = 7.84$, $p < .05$. Deviations from linearity were not significant ($p < .05$). This trend toward linearity gives a weak indication of the individual stroke's being the unit of programming in this context. However, there was no linear trend for RT in Experiment 3 $F(1, 9) = 1.74$. Since all of the line-segment tasks in these experiments consisted of connected strokes, there is some degree of "co-ography" here, more so in Experiment 3 than in Experiment 2. One suspects that the effective linearity may strengthen with isolated, iterative strokes.

Comparing the items in Experiment 2 with those in Experiment 3, the RTs were, on average, 21.6 msec faster for the former, with differences ranging from 14 to 33 msec. The differences for lengths 1-5 are all highly significant ($p < .002$). There is also a significant effect of items for both experiments ($p < .002$). This provides some indication that subjects treated line segments and curves differently.

If the unit of programming in handwriting is a complete cycle at one phase setting of agonist antagonist muscles, as argued by the spring model,¹ then the latency

functions should increase from stroke 2 to stroke 4 in Experiment 2 and from strokes 2 to 4 to 6 in Experiment 3, since these are the items with complete phase cycles. A significant increase was found for Experiment 2 ($p < .002$); however, no appreciable RT difference was detected for Experiment 3. Thus, phase cycles could be a significant factor in programming connected, alternating straight-line segments, but do not seem to be a determinant in planning handwriting sequences with continuous curved lines.

The letter has also been advanced as a possible unit of programming. Hulstijn and van Galen¹⁵ found a linear increase in RT with a sequence length of 1 to 4 letters, although the slope of this latency function decreased rapidly with practice. Considering that *v* and *w* are each single letters, if the unit were this large, one would expect a flat latency function. The difference in the mean latencies is significant, however ($p < .002$). The same finding holds true for comparison of the letters *u* and *ω* in Experiment 3 ($p < .01$). Thus, it would seem that in this context the letter was not an integrated unit.

Another possibility to be considered here, and one that seems most promising, is that the unit by which the program for handwriting output is built up is the downstroke. Taking into account just the number of downstrokes in Experiment 2, we find that RT increases from 196 msec for one downstroke to 198 msec for two to 204 msec for three. Similarly, in Experiment 3, RT increases from 209 msec for *u* to 218 msec for *ω*. An efficient way to mentally or verbally code the task sequences is by using a form of rhythmic notation, that is, "1-and" = *v*; "1-and-2" = *u*, *u*; "1-and-2-and" = *w*, *u*; "1-and-2-and-3" = *W*. An analogy can be drawn here to the idea of Sternberg *et al.* that the stress group is a unit of programming in speech where a primary stress can be followed by one or two unstressed syllables or words to form a unit.

Movement Time

FIGURE 4 shows the mean duration or total movement times (that is, the time from the first movement of the pen until the end of the last stroke) for each sequence length in Experiments 2 and 3. Data from both experiments are well described by linear functions, indicating that mean duration increased approximately linearly with the length of the sequence. Using a least-square method for linear regression, a slope of 112 msec was obtained in Experiment 2, while a slope of 95 msec was obtained in Experiment 3. The quadratic component predicted by the model of Sternberg *et al.*¹⁷ was not evident here. Rather, each stroke added a constant amount of time to the total duration. However, the slopes of the fitted linear functions differ significantly between experiments, with durations for the items in Experiment 2 being 18 msec slower on average. Thus, although the writing of these items is initiated reliably faster, it is carried out more slowly. Worthy of mention is the highly significant ($p < .002$) difference between the means for $n = 3$ and $n = 5$ in the two experiments. It can be seen that the items written more slowly are those in Experiment 2, which end in a downstroke with an incomplete phase resolution. This suggests that a similar mechanism may be effective for RT and MT and contributes further evidence that the phase cycle and the downstroke may delimit a unit of programming in handwriting. It may be argued that the items for $n = 3$ and $n = 5$ in Experiment 3 also end in a downstroke with an incomplete phase cycle, but show no significant departure from linearity. One possible explanation is that subjects indeed treated the items in the two experiments differently, with the downstroke of a curved segment not being as well defined. Admittedly, this interpretation is speculative.

Discussion

The results of these experiments seem to indicate that the unit of programming in handwriting may be smaller than the letter, at least in certain contexts. There is some evidence that the downstroke may be a significant factor in determining the boundaries of a unit. The number of complete phase cycles may play a role in distinguishing units in Experiment 2, but there is clearly no difference in RT among items with differing numbers of complete phase cycles in Experiment 3. It may be that, in the case of temporally overlapping curved segments, once a particular phase is set any number of cycles may be run off. In other words, there are no differential effects of interactive phases evident in the reaction-time interval. Considering the MT data, items with complete phase cycles are more alike in the two experiments than the other items, indicating a possible significance of this factor.

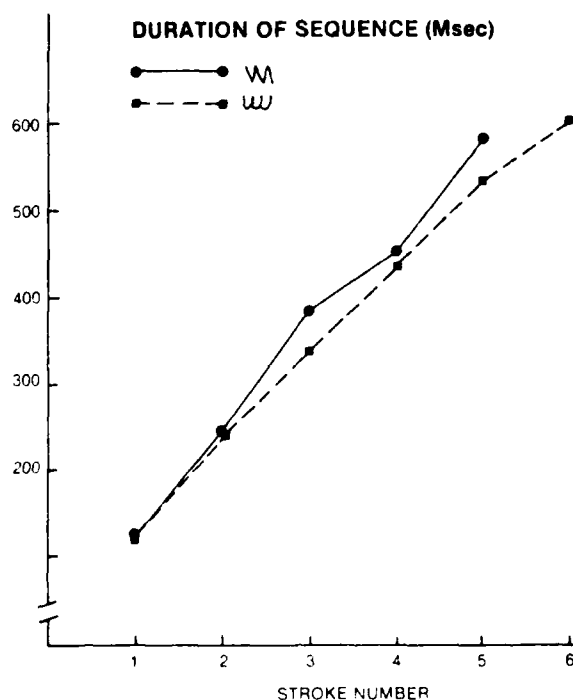


FIGURE 4. Duration of sequence in msec for each of the line and line-curve segment experiments plotted as a function of stroke number

GENERAL DISCUSSION

The main result obtained in the first experiment, namely, an effect of reducing the opportunity for advance planning, upon the stroking characteristics of the first letter of the sequence, seems to indicate that subjects began to write before they were fully prepared, then adjusted their motor program on-line. In this "reparameterization" condition they may not have taken time to complete programming prior to movement; therefore, the process of adjusting the program may have lingered over into the execution of the first stroke. Thus, size prescription (force-time relationship) of a handwriting sequence seems to be an integral aspect of a motor program.

The notion of space-time invariance¹⁷ is supported by similarities in stroke parameters for length and velocity. The shape of the duration curve also indicates the

presence of an invariant timing characteristic. Thus, the ability to program or plan a handwriting sequence in advance can be seen to affect the RT interval, and to carry over into the characteristics of the movement itself, while the ratios among the strokes remain relatively intact. Time and space, letter-size in this case, are then not necessarily consequences of each other, but may be general organizational principles underlying handwriting.

Information about the elements of organization in handwriting is addressed by focusing on the results of the last two experiments. Tentative evidence is provided for the downstroke, possibly integrated in some way with the phase cycle, to be a unit of programming in these handwriting tasks. Incidental support is found for Sternberg *et al.*'s²⁰ idea of a length-dependent increase in RT, although a quadratic component for MT was not evident. The general indication seems to be in agreement with Keele²¹ and Wing²¹ that timing is not only an integral part of motor sequencing, but also aids in the specification of programming units. It has been demonstrated in these experiments that some aspect of relative timing is invariant in the stroking characteristics for a handwriting sequence and that particular details of timing within the sequence, such as rhythmical patterning perhaps, are related to the preparatory time interval in terms of unit specification and total time.

REFERENCES

1. KEELE, S. W. 1981. Behavioral analysis of movement. *In* Handbook of Physiology, Vol. II Motor Control, part 2. V. B. Brooks, Ed. American Physiological Society, Baltimore, MD.
2. STELMACH, G. E. & V. A. DIGGLES. 1982. Control theories in motor behavior. *Acta Psychol.* **50**: 83-105.
3. HENRY, F. M. & D. E. ROGERS. 1960. Increased response latency for complicated movements and a "memory drum" theory of neuromotor reaction. *Res. Quart.* **31**: 448-458.
4. KEELE, S. W. 1968. Movement control in skilled motor performance. *Psychol. Bull.* **70**: 387-403.
5. MERTON, P. A. 1972. How we control the contraction of our muscles. *Sci. Am.* **226**: 30-37.
6. RAIBERT, M. H. 1977. Motor control and learning by the state-space model. Technical report AI-TR-439. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.
7. STERNBERG, S., S. MONSELL, R. L. KNOLL & C. E. WRIGHT. 1978. The latency and duration of rapid movement sequences: Comparison of speech and typewriting. *In* Information Processing in Motor Control and Learning, G. E. Stelmach, Ed. 118-150. Academic Press, New York, NY.
8. STERNBERG, S., C. E. WRIGHT, R. L. KNOLL & S. MONSELL. 1980. Motor programming and rapid speech: Additional evidence. *In* Perception and Production of Fluent Speech, R. Cole, Ed. 508-533. Lawrence Erlbaum, Hillsdale, NJ.
9. SHALTER, L. H. 1980. Analysing piano performance: A study of concert pianists. *In* Tutorials in Motor Behavior, G. E. Stelmach, Ed. 443-456. North-Holland, Amsterdam.
10. SHALTER, L. H. 1978. Timing in the motor programming of typing. *Quart. J. Exp. Psychol.* **30**: 333-345.
11. VREDENBREGT, J. & W. G. KOSTER. 1971. Analysis and synthesis of handwriting. *Philips Tech. Rev.* **32**: 73-78.
12. ELLIS, A. W. 1982. Spelling and writing. *In* Normality and Pathology in Cognitive Functions, A. W. Ellis, Ed. Academic Press, London.
13. HOLLERBACH, J. M. 1981. An oscillation theory of handwriting. *Biol. Cybernet.* **39**: 139-156.

14. VAN GALEN, G. P. 1980. Handwriting and drawing: A two-stage model of complex motor behavior. *In* *Tutorials in Motor Behavior*, G. E. Stelmach & J. Requin, Eds.: 567-578. North-Holland, Amsterdam.
15. HUISTIJN, W. & G. P. VAN GALEN. 1983. Programming in handwriting: Reaction time and movement time as a function of sequence length. *Acta Psychol.* **54**: in press.
16. TEULINGS, J. L. H. M. & A. J. W. M. THOMASSEN. 1979. Computer-aided analysis of handwriting movements. *Visible Lang.* **13**: 219-231.
17. VIVIANI, P. & V. TERZOLO. 1980. Space-time invariance in learned motor skills. *In* *Tutorials in Motor Behavior*, G. E. Stelmach & J. Requin, Eds.: 525-539. North-Holland, Amsterdam.
18. LA BERGE, D. H., P. VAN GELDER & J. YELLOTT. 1980. A cueing technique in choice reaction time. *Percept. Psychophys.*: 51-67.
19. LARISH, D. & G. E. STELMACH. 1982. Preprogramming, programming and reprogramming of aimed hand movements as a function of age. *J. Motor Behav.* **14**: 322-340.
20. STELMACH, G. E. & J. L. H. M. TEULINGS. 1983. Motor programming and temporal characteristics in handwriting. *Acta Psychol.* **54**: in press.
21. DENIR VAN DER GON, J. J. & J. T. THURING. 1965. The guiding of human writing movements. *Kybernetik* **2**: 145-148.
22. BONNET, M., J. REQUIN & G. E. STELMACH. 1982. Specification of direction and extent in motor programming. *Bull. Psychonom. Soc.* **19**: 31-34.
23. WING, A. M. 1978. Response timing in handwriting. *In* *Information Processing in Motor Control and Learning*, G. E. Stelmach, Ed.: 153-172. Academic Press, New York, NY.

The Patterning of Time and Its Effects on Perceiving

MARI RIESS JONES

*Department of Psychology
The Ohio State University
Columbus, Ohio 43210*

The problem addressed here involves the patterning of events in time and the way this patterning influences a person's attention to and perception of pitch relationships that are embedded within simple music-like tonal sequences. This paper reports the outcome of the first of several experiments on this problem that I have been conducting at Ohio State with the invaluable assistance of Beth Marshburn and Gary Kidd. These studies all consider the special effects that might arise from some larger rhythmic context, one that is set up within the experimental session itself, upon a person's ability to recognize melodic relationships within individual auditory patterns.

The general idea involves manipulation of the rhythmic context afforded by certain patterns and pairs of patterns in a melodic recognition task. A melody recognition task is a task wherein, on any given trial, in a multiple-trial experimental session, a person hears two auditory patterns, the first one being a standard pattern and the second being a comparison pattern. The listener's task is to decide whether the second melody is the same or different from its standard, when, in fact, half of the time it contains some melodic changes of interest. In the present case we are interested in a change involving the pitch of a single note in the comparison pattern. Effectively this changes two pitch intervals in that pattern, where pitch intervals refers to the pitch distance between a pair of notes.

FIGURE 1 shows a typical recognition trial in one of our studies: a listener hears a warning tone; then, after a pause, the standard pattern of a pair occurs, and it is followed then by the comparison pattern; finally, a response period of 6 seconds is presented. The patterns are temporal patterns, each involving 12 square-wave tones with fundamental frequencies corresponding to tones of the C major scale. FIGURE 1 also shows a typical standard and comparison melody based on these notes. Comparison melodies, if they differed from the standard, would always preserve its basic shape and the changed note would still be consistent with the key of C major. In this example, a changed comparison pattern is shown: the pitch of the note occurring at the 11th serial position is shifted up one scale step, changing its relative distance from immediately neighboring notes. A deviation, if present, could occur at one of three different serial positions.

We created many such melodies using some systematic rules for arranging pitch intervals and contour. And since all patterns were necessarily temporal in nature, the melodies so constructed also had a rhythm. Indeed, rhythm was the variable of major interest here. The pair of melodies presented on any given trial always shared the same rhythm, but rhythm could differ between pairs of patterns encountered on different trials in an experimental session. A given pair of patterns was developed either in an isochronous rhythm or in what is termed here a duple rhythm. These two rhythms are shown in FIGURE 2. The isochronous rhythm is based on a recurrent stimulus-onset-asynchrony (SOA) of 300 msec, whereas the duple rhythm consists of a long-short time pattern with alternating SOAs of 300 msec and 200 msec, respec-

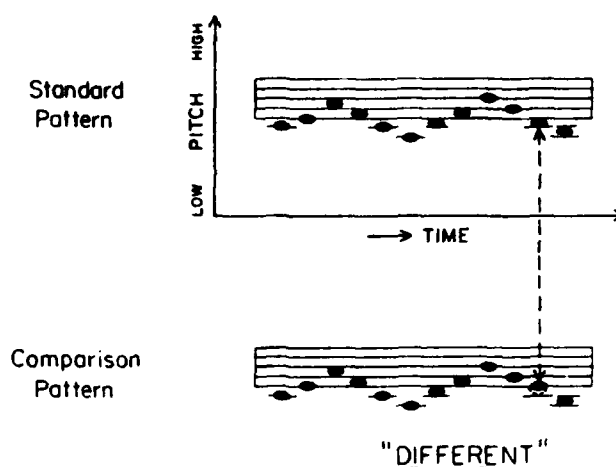


FIGURE 1. A single trial presented a listener with a standard melodic pattern (*top*) and then a comparison (*bottom*) when the comparison contained (half of the time) a changed tone (*broken line*).

tively. (In both rhythms, these SOAs reflect primarily tonal durations; off-times between tones were constant at 5 msec.) Obviously, the difference between the two rhythms is that the duple rhythm introduces *agogic* (that is, time-based) accents through relatively lengthened events. These accents fall on the first tone in a pattern and on all odd-numbered serial locations. Musically speaking, it is possible to indicate the duple rhythm as $\bullet \cdot \bullet$ and the isochronous as $\bullet \cdot \bullet \cdot$ or $\bullet \cdot \bullet \cdot \bullet \cdot$, and so forth, where $\bullet = 200$ msec. Both rhythms, however, hold constant (at 300 msec) the duration of a changed note (at serial positions 5, 7, and 11, shown by arrows in FIGURE 2) when such occurs in a comparison pattern.

Our interests in the effects of rhythm upon melody recognition have arisen for several reasons. One reason involves purely experimental considerations. In a number of earlier studies we uncovered some intriguing effects of rhythmic context upon judgments about temporal aspects of pattern structure such as temporal order and tone

Two Rhythms

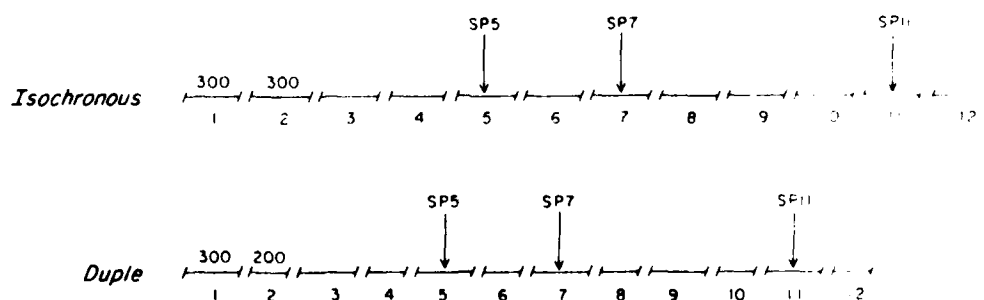


FIGURE 2. Two rhythms used in the present study: The isochronous rhythm (*top*) has an SOA of 300 msec throughout, while the duple (*bottom*) has alternation of SOAs of 300 msec and 200 msec.

duration. In these studies, there was some suggestion that when rhythm is varied as a within-subject so that a listener encounters several different rhythms within a single experimental session, overall performance is lowered relative to the case where rhythm is treated as a between-subject variable.^{1,2} In the present series of studies we extend this investigation to the examination of recognition judgments about the nontemporal structure of melodies, namely to perception of pitch relationships embedded within a melody.

Our interest in choosing to extend this investigation specifically to perception of melodic relationships highlights another motivation for this research. There is also a theoretical reason for this inquiry. Briefly, this relates to ideas, which I have developed elsewhere, that attending to dynamic patterns is itself a dynamic activity responsive to relative time relationships within unfolding patterns.^{3,4} This means that the rhythms of patterns encountered by a listener can guide attending so as to prepare the listener for "when" certain future serial events will occur. In this view, perceiving the musical relationship between a pair of tones will be more accurate if attentional energy is targeted to the temporal occurrence of this tonal interval. The idea is that a person abstracts rhythmic invariants and uses these to guide attending. This, in turn, immediately raises the issue of the internal "representation" of rhythm. Without belaboring this issue, let me simply state that I find it most parsimonious to subscribe to a biological model and to merely assume that all organisms are essentially rhythmic. Others who have considered the timing and coordination of motor productions have made similar assumptions.^{5,6} From the perspective of an attentional theory, this means that when one encounters temporally patterned energy in the environment, one is capable of responding in kind with one's own physical, attentional energy in a sort of "act of synchrony." Through this act, internal organismic rhythmicities are awakened, and I assume that these help to guide future attending over time. Indeed, when we consider that rhythm is defined in terms of relationships between time periods, it is not unreasonable to assume that abstracted rhythmic invariants guide attending over several time periods simultaneously where relatively smaller time intervals are contained in (nested) relatively longer ones.^{3,7}

In the present study, the rhythms involved can be described in terms of invariant time ratios that obtain between neighboring and remote events. For example, the ratio of SOAs between neighboring events in the isochronous rhythm is 1/1 whereas for the duple rhythm it is 3/2. Such ratios specify temporal grouping (or lack of it). That is, because the isochronous rhythm has a ratio of 1/1, for a sequence of 12 tones, it lends itself to several potential groupings with accents on every other note, for example, or on every third, or every fourth note. The duple rhythm, however, tends to suggest lower-order groups of two notes with the first element accented. At the higher temporal level, both rhythms permit equal spacing of accents in time, although the actual periods between accents differ for the two rhythmic types. For example, even if both patterns are grouped in twos, the higher-order time period operating between accents in the isochronous rhythm spans 600 msec, whereas that of the duple spans 500 msec.

Thus, if temporal contexts can direct attending "toward" or "away from" certain temporal locations, this may enhance or retard pickup of related melodic information. Targets of temporally guided attending are accents. By manipulating the rhythmic context to induce accenting within individual melodic sequences, we recently found evidence to support this prediction.⁸ In this study, we held constant the melodic and temporal relationships between three central tones within nine-tone auditory sequences while varying the timing of the surrounding tones. When contextual timing induced attentional targeting to the temporal location of a possible pitch change in the central tone group of a comparison pattern, persons were much more accurate at identifying it as different from the standard.

In the present study we go farther and manipulate the rhythmic environment around a pair of melodies. From the data afforded by the study of Jones, Boltz, and Kidd⁸ discussed above, we determined that often the regularity of isochronous rhythms is as effective as certain agogic-accent patterns in directing attending to critical serial positions within a sequence of events. Thus, in the present study we anticipated that the two rhythms we selected might be of comparable effectiveness in guiding attending within a pattern or over a pair of patterns sharing that rhythm on any given trial, all other things equal. But we are interested in effects that might emerge over trials if these two rhythms were combined in a single session. Is it possible that the temporal structure of surrounding pattern pairs (that is, those occurring on other trials) influences the course of attending locally within sequences of a given standard-comparison pair? Since accent locations follow different respective trajectories in the isochronous and duple rhythms, when both occur in a single session could rhythmic carryover effects result that might establish an occasional inappropriate dynamic set? If so, a listener would target attending to the wrong temporal location and thereby miss a critical pitch interval in the standard or comparison melody. Thus, experimentally, the question of interest becomes: Does it matter if persons experience only one rhythm within an experimental session or if they experience two?

The experimental design addressed to answering this question is outlined in TABLE 1. Subjects participated in 24 trials on one of the two rhythms (isochronous or duple) in a rhythmic adaptation phase and then, without a break, they participated in a test phase consisting of a total of 84 more trials. In the test phase, two kinds of trials

TABLE 1. Experimental Design

Adaptation Phase	Test Phase	
	Context Trials	Test Trials
Isochronous	Isochronous	Isochronous
Rhythm	Isochronous	Duple
Duple	Duple	Isochronous
Rhythm	Duple	Duple

occurred: (1) context trials, which occurred in runs of two and three (equally often), and in which melodic pairs always preserved the adaptation rhythm; and (2) test trials, which always occurred singly and in which the two melodies of a pair *either* both shared the adaptation rhythm or they both were switched to a new test rhythm. The basic idea is simple: in the adaptation phase persons get used to anticipating future events on the basis of rhythmic invariants in a given adaptation rhythm. This rhythm continues into the test phase as the context rhythm. If in the test phase this context rhythm is supplemented on test trials with the same rhythm, then performance here will be relatively good. On the other hand, if in the test phase a person encounters a novel rhythm in addition to the context rhythm, there now is some basis for interference. On some trials attentional rhythmicities excited by one rhythm will inappropriately carry over to the next trial and thereby reduce the effectiveness of temporal targeting of attention. In terms of the relative timing of points where deviant pitch relationships could occur, the two rhythms differ and so abstraction of the rhythmic invariants of one should inappropriately guide attending in the other. Thus, when text and context rhythmicities differ, performance in melody recognition should suffer.

To summarize, two different rhythms were used to study the effects of larger

temporal context in inducing attentional targeting over time. In an adaptation phase, where listeners experienced only one of the rhythms, large differences in melodic recognition were not expected. In a test phase where listeners could, in some conditions, experience a combination of both rhythms, melody recognition performance was expected to be poorer than in the case when listeners encountered the same rhythm on all trials. In short, in the test phase we predicted a context rhythm by test rhythm interaction wherein performance was poorer whenever context and test rhythms differ.

METHOD

Subjects

Sixty subjects participated in the experiment in return for credit in an introductory psychology course. Twelve subjects were eliminated on the basis of failure to differ significantly from chance performance during adaptation. Questionnaires indicated that subjects had a background of musical study ranging from 0 years to 13 years. The median number of years of music study was 3.5 years for the 48 individuals whose data were analyzed.

Apparatus

All tones were square waves generated by a Wavetek Model 159 waveform generator controlled by a Cromemco Z-2 microcomputer. A 20-msec rise and a 20-msec fall time was imposed on each tone. Stimulus sequences were tape-recorded on a Nakamichi LX-3 cassette recorder. Prior to recording, all tones were equated for subjective loudness according to criteria of three judges.

The prerecorded tapes were played on a Tandberg Model TCD 310 MK II cassette recorder at a comfortable listening level over AKG K-240 headphones.

Materials

Melodic manipulations involved a total of 18 standard patterns. Each pattern consisted of 12 tones arranged serially according to rules applied to the C major diatonic scale. Patterns were all constructed using the following guidelines: Regardless of a pattern's initial tone (which was randomly selected), the scale distances between note pairs spanning serial locations 1-2, 3-4, 5-6, 7-8, 9-10, and 11-12 were constant within a pattern. These distances could realize either ± 1 , ± 2 , or ± 3 scale steps (that is, next N^j rules where $j = 1, \pm 2, \pm 3$). This constraint provided a basis for grouping notes into pairs within a sequence. Contour was also systematically varied so as to provide an alternative basis for grouping involving clusters of three tones having a common pitch trajectory (for example $+$, $+$, $+$ or $-$, $-$, $-$). This was realized through constraints imposed upon the pitch interval connecting successive pairs of tones within a sequence: the direction of this interval, whether up ($+$) or down ($-$) in pitch, was selected so as to induce sets of at least three ascending or descending tones. Thus, pattern contours could be of the form $+$, $+$, $-$, $-$, $+$, $+$, $-$, $-$, $+$, $+$, $-$, $-$ or $-$, $-$, $+$, $+$, $-$, $-$, $+$, $+$, $-$, $-$, $+$, $+$, $-$, $-$, and so forth. The result of these two kinds of rules was a large set of melodies that afforded grouping into pairs or into triplets of tones. The underlying

rationale was to build a set of melodies potentially compatible with grouping properties of a range of rhythms. From this set a number of resulting melodies were eliminated using the following criteria: (1) Melodies containing two temporally adjacent occurrences of a given tone; (2) melodies containing successive tone groups that were either identical or were exact reversals (for example, A₄ G₄ E₄ G₄ A₄ within a sequence was not allowed); (3) melodies with notes below G₃ or above G₅.

From each of the 18 remaining standard melodies, four types of comparison melodies were constructed: (1) same comparison in which the comparison pattern was identical to the standard; (2) deviant comparison with a changed note at the fifth serial position (SP5); (3) deviant comparison with a changed note at the seventh serial position (SP7); (4) deviant comparison with a changed note at the eleventh serial position (SP11). Any deviant note, if it occurred, never occurred at an extreme point (that is, highest of lowest note of pattern) in the pattern and it was always an in-key contour-preserving note change of either ± 1 or ± 2 semitones (ST) (with the majority involving a 2-ST shift).

Pairs of standard and comparison patterns were constructed and arranged in two different random orders. In each order, half of the time the unchanged comparison patterns were paired with the standard and half of the time a deviant comparison was paired with the standard (with deviant comparisons involving SP4, SP7, and SP11 occurring equally often). The ordering of pattern pairs was constrained according to experimental design requirements that a given session contain an adaptation phase, in which pattern pairs occurred only in a context rhythm, and a test phase, in which pattern pairs could appear in either the pre-established context rhythm or in a test rhythm. The adaptation phase consisted of 24 trials of randomly ordered pattern pairs arising from four different standard melodies. The test phase consisted of 84 trials of which 60 were termed context-rhythm trials and 24 were termed test-rhythm trials. The constraint in the test phase was that a single test-rhythm trial was programmed to occur (equally often) after either two or three context-rhythm trials. Within both the context-rhythm and test-rhythm trials sets, equal numbers of "same" and "different" standard comparison pairs occurred. Melodies in the test phase were based on the remaining 14 standard patterns not used in the adaptation phase.

Rhythmic manipulations involved imposing upon these melodies one of two rhythms, depending upon the experimental conditions. The rhythms, isochronous and duple, are shown in TABLE 1. The isochronous rhythm was based upon stimulus-onset asynchrony (SOA) of 300 msec; the duple rhythm was a long-short recurrent pattern, wherein the 300-msec tone was now relatively long and the 200-msec tone was shorter, thus placing agogic accents on serial positions 1, 3, 5, 7, 9, and 11, assuming that the longer events would be heard as accented. (This was, in fact, the case according to several independent judges.)

Design

The design consisted of an adaptation phase and a test phase and data from the two phases were analyzed separately. In the adaptation phase, context rhythm (isochronous, duple) and test rhythm (isochronous, duple) conditions yielded a 2×2 between-subjects' factorial design. In the test phase, a $2 \times 2 \times 2$ factorial design involved context rhythm (isochronous, duple), test rhythm (isochronous, duple) and trial type (context, test). Only trial type was a within-subjects' variable. Twelve subjects were randomly assigned to each of the four context $\times$ test conditions.

Procedure

Subjects were instructed, via prerecorded tapes, that they had to judge whether the second melody of a pair contained a single changed note (they received an illustrative diagram) on each of a series of trials. They were told that half the time the second melody would, in fact, be different from the first and furthermore that a change would never violate the contour, or shape, of the melody, only the pitch distance. They responded to each pair in writing by checking "same" or "different," plus indicating certainty in their judgment, using a score of 1-5 (1 = very certain, 5 = uncertain).

Participants listened to the melodies over AKG (Model K-240) headphones and responded to a total of 108 trials plus two practice trials. A high-pitched 1-sec warning tone (3,500 Hz) signaled the onset of each trial. This was followed by a 2-sec silence and then by the pair of patterns separated from one another by a 3-sec pause. A 6-sec response interval intervened between the offset of a comparison pattern and the onset of a warning tone signaling the next trial.

Subjects were tested in groups of two to four persons, but they were individually separated from one another by partitions, within a sound-deadened room.

RESULTS

Scoring

Both PC , proportion correct, and A_g , a nonparametric ROC measure of recognition accuracy, were computed for each subject. The A_g score is computed by the trapezoidal rule^{9,10} based on a program developed by Davison and Jagacinski.¹¹ An A_g score estimates unbiased recognition accuracy in the two-choice case where chance is .50. An A_g score of 1.00 reflects perfect discrimination and .50 is random guessing. Analyses were carried out on both dependent measures with results essentially equivalent for both. For this reason, only the results of A_g are reported here.

Adaptation Phase

The four different groups of subjects who were assigned to context-test combinations in the test phase were compared with respect to their performance on the first 24 adaptation trials. No significant differences among the four groups was observed $F(1, 44) < 1$ for the context-by-test interaction, indicating that these subjects were performing indistinguishably from one another prior to the test phase. It is also important to note that melody recognition accuracy was not significantly affected by the context rhythm (isochronous, duple) in adaptation. These rhythms were equally effective in determining recognition accuracy, $F(1, 44) < 1.00$.

Test Phase

Means and standard deviations for A_g scores as a function of context-test combinations for these two trial types are in TABLES 2 and 3. In all conditions performance was significantly above chance levels on this measure ($p < .05$). The data in TABLES 2 and 3 show that in both test and context trials subjects were better in those conditions where the rhythms remained the same across context and test trials than

TABLE 2. Means and Standard Deviations (A_g) on Context Trials in Test Phase

Context Rhythm	Test Rhythm	
	Isochronous	Duple
<i>Isochronous</i>		
Mean	.68	.60
S.D.	.10	.10
<i>Duple</i>		
Mean	.59	.66
S.D.	.08	.09

when the rhythms differed. Overall, the context-by-text rhythm interaction was statistically significant, $F(1, 44) = 4.40, p < .05$. A significant three-way interaction of trial type with context-by-text rhythm did not emerge, indicating that introduction of a novel test rhythm during the test phase has a pervasive effect. Subjects are not merely unable to lock into this new rhythm, but there is also a pronounced "boomerang" effect which lowers performance with the contextual, or adapted-to, rhythm when a novel test rhythm is suddenly inserted into the experimental session.

DISCUSSION

The data are consistent with the hypothesis that the larger rhythmic context within an experiment can influence attending to serial relationships that are not primarily temporal. Poorer performance observed in the test phase for listeners who experienced two rhythms suggests the importance of temporal patterning in guiding attending.

However, even in conditions where performance is best, namely, in the unchanged rhythm conditions, melody recognition does not appear, at least superficially, to be exceptionally good (that is, A_g scores of .66-.72). But this is deceptive. If we weigh this performance against what is known about melody recognition in comparable situations, this study actually reflects fairly good performance. That is, the pattern recognition task was difficult for at least three reasons: (a) all pitch interval deviations reflected in-key violations; and (b) all pitch interval deviations did not break contour; and (c) all patterns were relatively long (12 tones). The work of Lola Cuddy^{12,13} and Jay Dowling^{14,16} has shown repeatedly that the first two features lead to very poor melody recognition. And my own research, as well as general expectations, suggests pattern recognition difficulty increases with pattern length. When these facts are considered, it turns out that performance in the present task is relatively good.

Thus, it appears that when fairly subtle pitch changes in a melody were introduced,

TABLE 3. Means and Standard Deviations (A_g) on Test Trials in Test Phase

Context Rhythm	Test Rhythm	
	Isochronous	Duple
<i>Isochronous</i>		
Mean	.64	.63
S.D.	.16	.13
<i>Duple</i>		
Mean	.66	.73
S.D.	.11	.11

it helps to be able to anticipate "when" in time these changes will transpire. I should mention that these findings are not encompassed by proposals that emerge from recent research on temporal coding such as Povel's¹⁷ or Handel's.¹⁸ This research is related to the present study in that it reveals the psychological salience of time ratios on performance in tapping tasks. But the formal descriptions of performance in tapping tasks are concerned primarily with the effects of *temporal* structure on *temporal* reproduction. Here, we find effects of *temporal* structure on *melody* recognition. The present data provide evidence for the general influence of higher-order temporal context upon attending and perceiving in auditory sequences.

There are some pragmatic implications of these findings. The most obvious one is that the decision to treat rhythm as a between- or within-subject variable should not be made lightly. Furthermore, we must weigh cautiously the generality of conclusions about listeners' performance levels in tasks that have exposed participants to a range of different rhythms and that have failed to take account of possible contextual effects.

Finally, consider some theoretical implications of these findings. One tempting means of conceiving of these effects at a more general level involves the idea of temporal uncertainty. That is, one could attempt to devise some metric that would reflect the fact that addition of a novel rhythm in certain conditions merely adds to the uncertainty about a pattern's rhythmic scheme. But is important to be careful about using this term "uncertainty" as a complete explanation of these effects. In the first place, in conditions with two rhythms there is, in fact, no uncertainty on many trials about *which* rhythm will occur on that trial because of the trial-to-trial contingencies of context and test trials. These trials were not randomly arranged and furthermore there were more context trials than test trials. And, of course, because standard and comparison melodies always share the same rhythm, if this contingency is considered, there is *no uncertainty* about the rhythm of a given comparison pattern. When these kinds of redundancies are considered, it quickly becomes clear that some information metric based on uncertainty will be a complex one subject to all the problems investigators encountered in the 1960s using these metrics to describe higher-order contingencies in pattern structure.¹⁹

There are other reasons why a simple rhythmic-uncertainty explanation seems incomplete. It is more descriptive than explanatory. Why should uncertainty in rhythmic structure affect sensitivity to melodic structure? Melodies remain the same across all four rhythm conditions. One might be tempted to rely here for an explanation upon some version of temporal parsing under uncertainty. However, it seems that when we finally become quite specific about what "parsing" means and how it is determined in this context, it is quite difficult to distinguish it experimentally from the simpler interpretation, which simply states that attending follows relative time rules that afford grouping. To be sure, uncertainty in temporal structure is in some way important to perception of nontemporal structure. My suggestion is that greater uncertainty, in a loose, descriptive sense, is detrimental because it affects the reliability with which the listener abstracts and uses appropriate rhythmic invariants to guide responding. On a number of trials, a listener may experience conflicting rhythmic expectations or may simply rely upon the wrong one.

REFERENCES

1. JONES, M. R., G. KIDD & R. WETZEL. 1981. Evidence for rhythmic attention. *J. Exp. Psychol.* 7: 1059-1073.
2. JONES, M. R. & M. BOETZ. 1981. Temporal context: What is it and what does it do for you? Paper presented at Psychonomic Society meeting, Philadelphia.

3. JONES, M. R. 1976. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychol. Rev.* **83**: 323-355.
4. JONES, M. R. 1981. A tutorial on some issues and methods in serial pattern research. *Percept. Psychophys.* **30**: 492-504.
5. KELSO, J. A. S., D. L. SOUTHARD & D. GOODMAN. 1979. The nature of human interlimb coordination. *Science* **203**: 1029-1031.
6. KELSO, J. A. S., B. TULLER & K. S. HARRIS. A 'dynamic pattern' perspective on the control and coordination of movement. *In* *The Production of Speech*. P. MacNeilage, Ed. Springer-Verlag. New York, NY.
7. YESTON, M. 1976. *The Stratification of Musical Rhythm*. Yale University Press. New Haven, CN.
8. JONES, M. R., M. BOLTZ & G. KIDD. 1982. Controlled attending as a function of melodic and temporal context. *Percept. Psychophys.* **32**: 211-218.
9. BAMBER, D. 1975. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J. Math. Psychol.* **12**: 387-415.
10. POLLACK, I., D. A. NORMAN & E. GALANTER. 1964. An efficient nonparametric analysis of recognition memory. *Psychonomic Sci.* **1**: 327-328.
11. DAVIDSON, T. C. B. & R. JAGACINSKI. 1977. Nonparametric analysis of signal detection confidence ratings. *Behav. Res. Methods Instrum.* **9**: 545-546.
12. CUDDY, L. L. & A. J. COHEN. 1976. Recognition of transposed melodic sequences. *Q. J. Exp. Psychol.* **28**: 255-270.
13. CUDDY, L. L., A. J. COHEN & D. J. K. MEWHORT. 1981. The perception of structure in short melodic sequences. *J. Exp. Psychol.* **7**: 869-883.
14. DOWLING, W. J. 1972. Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. *Percept. Psychophys.* **12**: 417-421.
15. DOWLING, W. J. 1978. Scale and contour: Two components of a theory of memory for melodies. *Psychol. Rev.* **85**: 341-354.
16. DOWLING, W. J. & D. S. FUJITANI. 1971. Contour, interval, and pitch recognition in memory for melodies. *J. Acoust. Soc. Am.* **49**: 524-531.
17. POVEL, D. J. 1981. Internal representation of simple temporal patterns. *J. Exp. Psychol.* **7**: 3-18.
18. HANDEL, S. & J. S. OSHINSKY. 1981. The meter of syncopated auditory polyrhythms. *Percept. Psychophys.* **30**: 1-9.
19. JONES, M. R. 1974. Cognitive representations of serial patterns. *In* *Human Information Processing: Tutorials in Performance Cognition*. B. Kantowitz, Ed. Erlbaum. Potomac, MD.

On Controlling Force and Time in Rhythmic Movement Sequences: The Effect of Stress Location

ANDRAS SEMJEN,^a ADELA GARCIA-COLERA, AND
JEAN REQUIN

*Department of Experimental Psychobiology
Institute of Neurophysiology and Psychophysiology
National Center for Scientific Research
Marseille, France*

INTRODUCTION

The work that we shall report has two facets: it relates to the rather broad problem of advance planning of rapid movement sequences; and it relates to the question of how the mechanisms that control timing and intensity of the component movements are coordinated in such rapid sequences.

The assumption that rapid movement sequences are planned in advance derives from at least two kinds of observations: those pertaining to structural properties of such sequences as they are executed, and those pertaining to the time it takes to prepare their execution.

The first approach is exemplified by Vince's experiment in which subjects drew vertical lines by continuously moving a pencil between two horizontal marks. When these movements were executed at a high speed, aiming errors were not corrected between individual movements, but rather between groups of movements (for example, between a first group in which all movements fell short of the mark, and a second group in which all movements attained the mark). From this observation the inference was drawn that, under the speeded-up condition of movement execution, decisions concerning movement extent were made for groups of movements, rather than for individual movements. Hence, the group was "planned" in advance as a whole.

The second approach is exemplified by studies from a number of investigators who, following Henry and Rogers' pioneer work,² attempted to relate the time cost of programming a movement sequence to its "complexity." Here the basic idea is that if time to initiate a rapid movement sequence (after a "go" signal) depends on the number of elements contained in the sequence, and/or on the specific relationships between these elements, this implies treatment of the sequence as a whole before its first element is emitted. "Treatment" may mean that a motor program is generated or activated for the execution of the whole sequence, much as Henry and Rogers' original memory-drum theory suggested. However, the minimum requirement for the advance planning notion is that the movement sequence be somehow identified as a particular response.

The reaction time (RT) analysis of how rapid movement sequences are planned and executed has been extensively applied to speech and typewriting.³ It has been

^aAddress for correspondence: C.N.R.S. I.N.P. 03, 31, Chemin Joseph-Aiguier, 13402 Marseille Cedex 9, France.

amply demonstrated that the RT of a spoken or written sequence increases as a function of the number of elements in the sequence (that is, the number of words or the number of keystrokes). Furthermore, it has been shown that this factor also affects the timing of the sequence, that is, the intervals between the component units.

The notion of advance planning of rapid movement sequences is clearly supported by the variations in RT provoked by the manipulation of the number of elements. However, the question can be raised as to whether more intrinsic properties of a sequence, such as the time distribution of motor intensities over the elements it contains, are also taken into account before the execution of the sequence starts. This question seemed to us of particular interest, since force, along with time, might constitute a basic dimension in the control of movements.

In a series of previous experiments, the results of which will be reported elsewhere, we asked subjects to produce rapid sequences of cadenced finger taps and to accentuate one of them. The sequence always contained the same number of taps (usually five), but the location of the stress varied within the sequence. The main finding from these experiments was that when the serial position of the element to be stressed was indicated to the subject by the go signal (a choice RT condition), the RT of the sequence was considerably longer than when the subject received such information prior to the go signal (a simple RT condition).

One possible interpretation of this "stress position uncertainty" effect is that stress location differentiates the whole sequence as a particular response. RT would then increase because, prior to its initiation, the required response must be identified, and perhaps programmed in accordance with the time and force distribution properties of the sequence. Since RT is known to increase as a function of the number of equally likely response alternatives,⁵ the obvious prediction from the above interpretation is the following: RT of rapid sequences of cadenced finger taps should increase as a function of the number of equally likely locations of the stressed tap within the sequence. The first aim of the experiment to be reported here was to test this hypothesis.

The second aim of the experiment was to elucidate the origin of the effect that stressing a tap had on the timing of the sequence. From our previous experiments it was established that stressing a tap lengthened the intervals that immediately preceded and followed the stress. While the variations in tapping cadence may reflect what Ostry⁴ calls "real-time organization of the sequence," they could also be caused by purely peripheral, biomechanical factors. For example, it can be argued that in order to produce stress, that is, more force, the subject moves his finger farther away from the key, and that the lengthened finger trajectory causes a lengthened tapping interval. While such an explanation could be valid for the interval that precedes the stress, it cannot hold true for the interval that follows it. Variations in the tapping cadence seem to depend on processes that control variations in the force level, rather than on processes linked to the force generation itself. In order to assess more carefully this possibility, we included in the present experiment a condition under which subjects were asked to tap one element of the sequence less forcefully than all the others. Even though under this "inverted stressing" condition the force patterns should be quite different from those observed under the ordinary stressing condition, we predicted that the variations in the tapping intervals before and after the critical ("negatively stressed") element would be approximately the same as under the ordinary stressing condition.

No explicit prediction was made as to the RT effect of "inverted stressing." However, since under this condition subjects were required to produce rather unusual force patterns, it was logical to expect longer RTs insofar as response organization concerned the sequence as a whole. Under the alternative hypothesis, namely, that only the very first element of the sequence is prepared during the RT period, the subject

would have to make the same binary choice between high or low level of force, independent of the stress condition. Hence the reaction time should remain constant.

METHOD

Task

Subjects were instructed to execute a sequence of four taps on a key with the index or middle finger of their preferred hand, reproducing a temporal model that was presented to them on each trial. This model consisted of four clicks of equal intensity, separated by 180-msec intervals. Subjects were also required to make one of the taps either stronger (stress + condition) or weaker (stress - condition) than the other three. These two conditions were run in separate series. The tapping response was to be started as quickly as possible after the display of a digit which indicated the serial position of the tap to be stressed in the ongoing trial.

The experiment took place in a noise-insulated chamber. The subject was seated facing the key, and a digital display unit. A trial began with a warning signal consisting of a "0" that was displayed for 500 msec. Two seconds later, the temporal model was delivered to the subject via headphones. After a 1-sec period, the display of a digit indicating the location of the stress signaled the onset of the movement sequence. The intertrial interval lasted 10 sec.

Measurements

All experimental events and measurements were automatically controlled by a PDP-12 computer (Digital Equipment Co.). The RT, measured as the delay between presentation of the digit used as response signal and onset of the first tap, and the duration of the intervals between the four taps were recorded for each trial. The force of each tap was measured in arbitrary units as the output voltage of a strain-gauge incorporated in the key.

A test was automatically performed on the measurements to determine whether a sequence had been executed as required. The trials in which the response was not correctly performed were repeated once at the end of each block. A response was classified as an error: (1) if the number of taps was other than four; (2) if the stress was not placed on the required tap; or (3) if the intervals exceeded certain limits that were fixed in advance. According to the temporal model, the second, third, and fourth elements of the movement sequence were to be tapped 180, 360, and 540 msec, respectively, after the first one. A time error was registered if any tap was executed more than 90 msec before or after its required delay.

Design

All subjects served in two experimental sessions. Half of the subjects received the stress + condition in the first session and the stress - condition in the second one, while the order was reversed for the other half.

The experimental session was divided into three parts. The purpose of the first part was to acquaint the subject with the task. It was designed as a simple RT situation in which the subject was instructed to stress a given tap before the beginning of a series of trials. Each possible stress location was practiced in a block of 10 trials.

The second and third parts of the session were conducted according to a choice RT paradigm. The subject was ignorant of the serial position of the stress until he was informed by the digit used as response signal. There was a four-choice RT condition and a 2-choice RT condition which were balanced across parts two and three. Under the four-choice RT condition, the four positions were presented as equally likely alternatives on all trials. This condition was run in three 32-trial blocks in which the four possible digits appeared in random order. Under the two-choice RT condition, there were only two positions in which the stress could be located. These two alternative positions were equiprobable. They were indicated to the subject before a series of trials began. All possible pairwise combinations between the four different stress locations were used. The six combinations (pairings) were tested in six blocks of 16 trials each. The order of presentation of these combinations was balanced across subjects.

Subject Selection

Twenty-four paid volunteers participated in a selection procedure which comprised the execution of five-tap sequences under simple and five-choice RT conditions. The task was performed in the same manner as the stress+ condition described above. The twelve participants with a higher proportion of correctly performed responses and faster RTs were selected as subjects for the experiment.

RESULTS

Errors

After its completion, each response sequence was tested for number of taps, stress location, and time intervals produced. Unlike testing for number of taps and stress location, testing for time intervals is arbitrary. No *a priori* definition can be given of the amount of discrepancy between the temporal model provided and actual response timing, which would unequivocally qualify a response as a time error. Nevertheless, we introduced this test in order to detect systematic departures from the required cadence and, most importantly, any strategy based on a partitioning of the sequence into well-separated subunits. Such a strategy on the part of the subjects would have compromised our attempt to utilize RT as an index of advance planning of the sequence as a whole.

The overall time-error rate was 5.9% and it was roughly the same under the stress+ and stress- conditions. Given the arbitrary character of the interval-testing procedure, individual mean scores for each of the response parameters recorded (that is, RT, interval duration, and tapping force) were computed in two ways: either excluding from or including in the computation the response sequences qualified as time errors. Manner of computation produced virtually no difference between group means for either RT or force. Interval mean and standard deviation were slightly different, depending on the way the computations were made. However, analyses of variance applied to these data showed the same factors to be significant (or nonsignificant), regardless of the manner in which individual scores were computed. Therefore, it was decided to include the "time errors" together with the correct responses in the presentation of the results.

TABLE 1 shows the combined error rates for number of taps and stress location, averaged over subjects. A greater number of errors was produced under the stress

TABLE 1. Average Error Rates

		Stress Location				Mean
		1	2	3	4	
Stress +	2 C	7.7	13.6	14.7	3.8	9.9
	4 C	13.6	14.8	16.9	11.8	14.2
Stress -	2 C	25.7	39.1	20.5	27.9	28.3
	4 C	26.0	32.4	24.4	22.6	26.3

than under the stress + condition. The effect of number of choices and stress location depended on the type of stress (+ or -). The number of choices caused a slight increase in error rate only under the stress + condition. Concerning the effect of stress location, the highest error rate was observed under the stress - condition when the stress was required on the second tap; the lowest error rate was observed under the stress + condition when the stress was required on the fourth tap. The comparison between error rate and RT data indicates no tendency for speed-accuracy trade-off.

Force Patterns

The force data are shown in FIGURE 1. Each point in the figure represents the mean force of a tap. The points have been grouped by four; these groups represent the four successive taps in a sequence. Stress locations are indicated by arrows. The force measurements recorded under the two-choice and four-choice conditions were pooled since there was no difference between the two conditions.

Stressed (either + or -) and nonstressed taps are sharply differentiated. This differentiation remains stable across all stress locations. However, the change in force

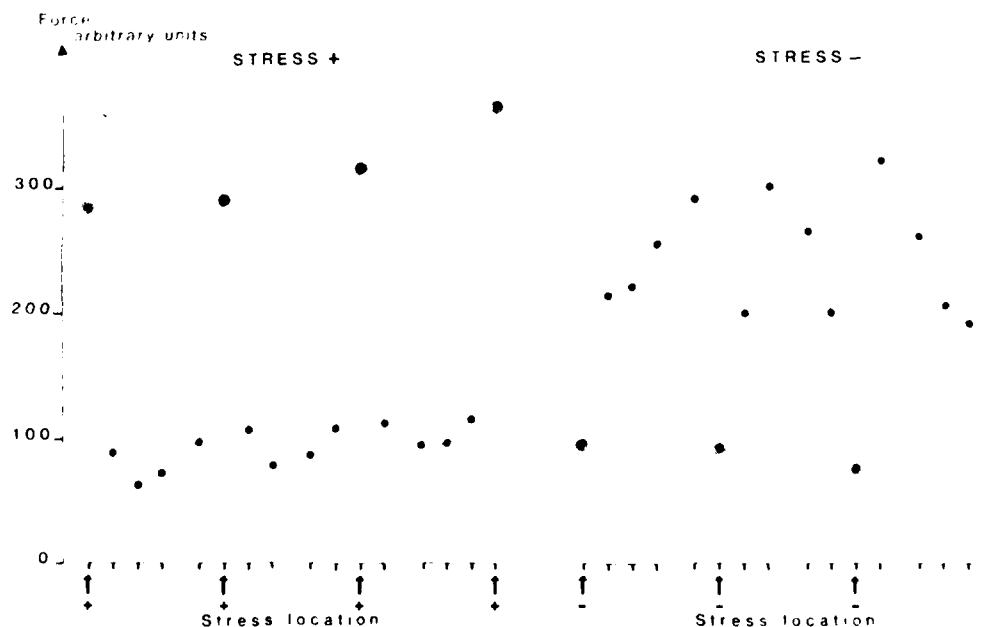


FIGURE 1. Mean force of the four successive taps in a sequence, plotted as a function of stress location (arrows) and stress type (+ and -).

level (from a nonstressed to a stressed tap and vice versa) seems to be more abrupt under the stress+ than under the stress- condition. As a consequence, the force level of the nonstressed taps presents much less variability under the former than under the latter condition.

The force of the stressed tap increases for stress+ and decreases for stress- as the stress moves from the first to the fourth position. Trend analysis confirmed that these effects were significant (for stress+: $F_{LIN} = 15.72$, d.f. = 1,11, $p < .01$ and $F_{QUA} = 6.72$, d.f. = 1,11, $p < .05$; for stress-: $F_{LIN} = 42.29$, d.f. = 1,11, $p < .001$).

The force level of the nonstressed taps shows a similar increase under the stress- condition ($F_{LIN} = 21.61$, d.f. = 1,11, $p < .001$), whereas it shows a more complicated, curvilinear evolution under the stress+ condition ($F_{QUA} = 19.99$, d.f. = 1,11, $p < .001$).

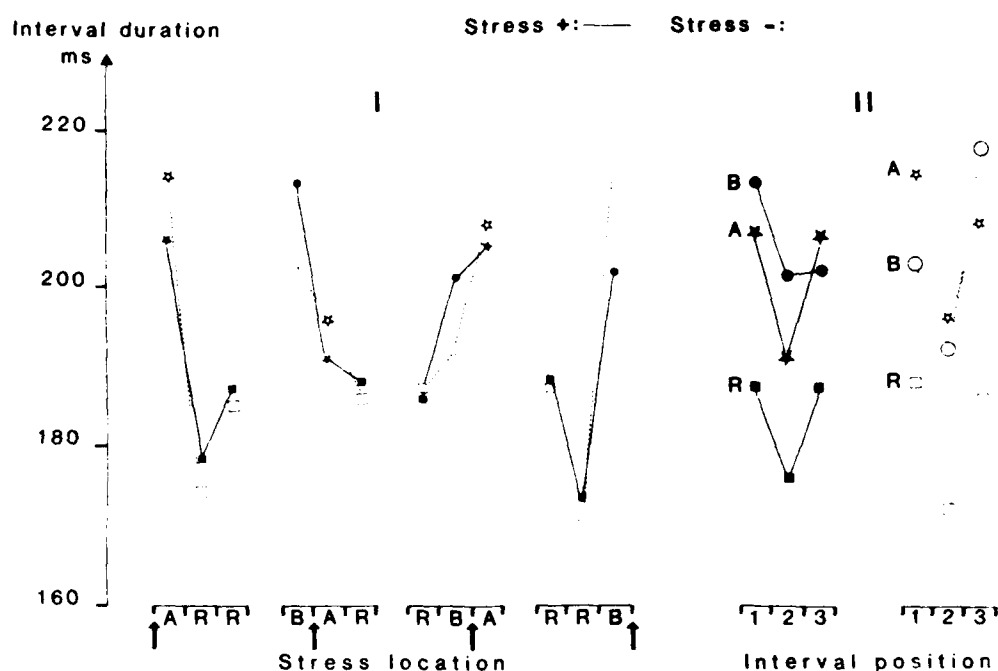


FIGURE 2. Part I (left): Mean duration of the three successive intervals in a sequence, plotted as a function of stress location (arrows) and stress type. Part II (right): Same data as in Part I, replotted according to interval position in the sequence (1, 2, 3), and interval position with respect to the stressed tap (B = before stress; A = after stress; R = remote). Remote intervals sharing the same position in the different sequences were averaged.

Sequence Timing

The mean duration of the successive intervals in a sequence (points grouped by three) is shown on the left (I) of FIGURE 2. Interval duration varied as a function of the interval's position (first, second, third) within the sequence and the location of the stressed tap. This was confirmed by a four-way analysis of variance (ANOVA), which yielded significant effects for stress location ($F = 10.89$, d.f. = 3,33, $p < .001$), interval position ($F = 24.33$, d.f. = 2,22, $p < .001$), and the interaction of these two factors ($F = 37.48$, d.f. = 6,66, $p < .001$). Neither the number of choices as a main factor, nor its interaction with the other factors, was significant. Stress type was not significant, but

the three-way interaction between interval position, stress location, and stress type reached significance ($F = 2.50$, d.f. = 6,66, $p < .05$). In other words, the joint effect of interval position and stress location was slightly modulated by stress type. However, the timing of the sequence was fundamentally the same for stress + and stress -, as can be seen in FIGURE 2 (I).

The time structure underlying all these interval configurations can be further clarified by regrouping the intervals according to their position in the sequence (first, second, third) on the one hand, and to their position with respect to the stressed tap (before, after, remote) on the other. Regrouping of the intervals according to this double classification leads to the time patterns shown on the right (II) of FIGURE 2. The intervals that are located before and after a stressed tap are longer than those located in other (remote) positions. In addition, the first and last intervals are longer than the second one.

Interval standard deviations (S.D.), averaged over subjects, are presented in FIGURE 3 (I). A four-way ANOVA yielded significant effects for stress location ($F = 6.14$, d.f. = 3,33, $p < .05$) and interval position ($F = 8.04$, d.f. = 2,22, $p < .01$), while number of choices and stress type were nonsignificant. The following interactions were significant: stress location by stress type ($F = 5.96$, d.f. = 3,33, $p < .05$), stress location by interval position ($F = 12.80$, d.f. = 6,66, $p < .001$), and stress location by interval position by stress type ($F = 7.94$, d.f. = 6,66, $p < .001$). The latter three-way interaction indicates that the effect of the stress on the interval that immediately precedes and follows the stressed tap strongly depends on the stress type.

On the right part (II) of FIGURE 3, interval S.D.s have been regrouped according to the double classification principle explained above. It can be observed that interval variability (that is, the regularity of timing) was more strongly influenced by "inverted stressing" than by ordinary stressing. It also appears that, unlike interval duration, interval variability does not present a unique configuration linked to the interval's position within the sequence.

One of the criteria for regrouping the intervals (on the right side of FIGURES 2 and

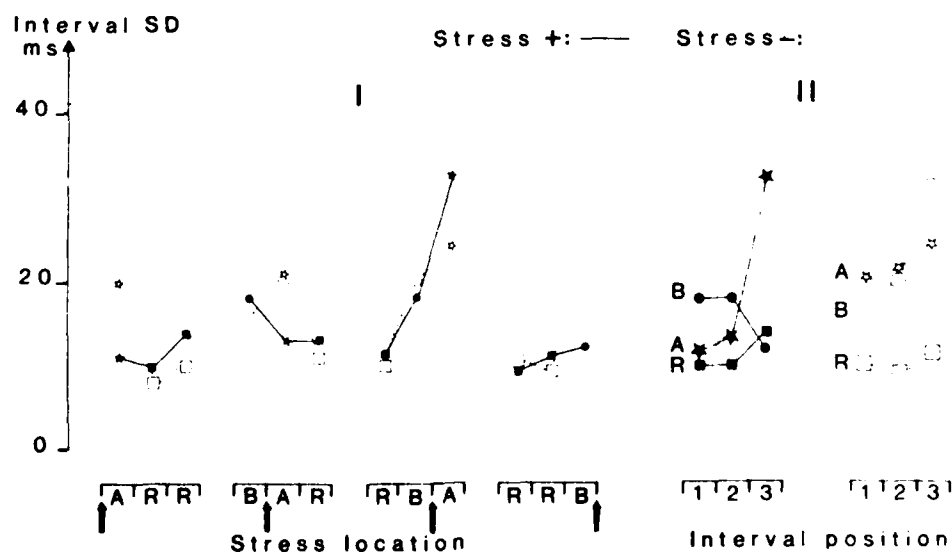


FIGURE 3. Part I (left): Mean standard deviations (SD) of the successive intervals in a sequence, as a function of stress location (arrows) and stress type. Part II (right): Same data as in Part I, replotted as in FIGURE 2, Part II.

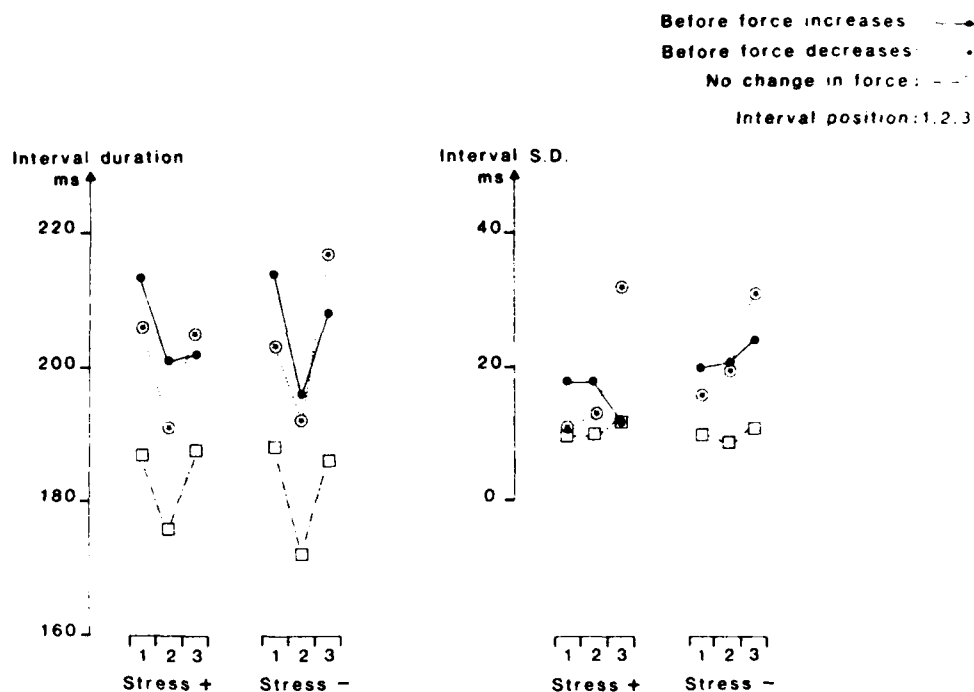


FIGURE 4. Interval mean duration and standard deviation (SD) plotted as a function of stress type (+ and -), interval position in the sequence (1, 2, 3), and interval position with respect to the force variations associated with stress production (before force increases, before force decreases, when no force change occurs). This figure is based on the same data as in FIGURES 2 and 3.

3) was their position with respect to stress location. It should be noted, however, that the direction of force change is opposite for the two types of stress. The interval produced before the stress precedes a force increase under the stress+ condition and a force decrease under the stress- condition. Conversely, the interval produced after the stress precedes a force decrease under the stress+ condition and a force increase under the stress- condition (see FIGURE 1).

In FIGURE 4, interval means and S.D.s are plotted according to the position of the interval in the sequence and with respect to the required variations in force level (before force increases, before force decreases, and when no force change occurs). The figure shows most clearly that variations in force level have the same lengthening effect on the duration of the intervals that precede these variations, regardless of the stress type. The interval data were subjected to a four-way ANOVA. Interval position within the sequence produced a significant main effect ($F = 25.08$, d.f. = 2,22, $p < .001$), whereas number of choices and stress type were nonsignificant. Position of the interval with respect to force change produced a significant main effect ($F = 29.91$, d.f. = 2,22, $p < .001$). Subsequent partial comparisons showed that intervals that preceded an increase and those that preceded a decrease in force did not differ significantly. However, the interaction of this position factor with interval position in the sequence was significant ($F = 9.41$, d.f. = 2,22, $p < .01$). This interaction signifies that, at the beginning of the sequence, a force increase had a greater lengthening effect on the interval than a force decrease, whereas at the end of the sequence this relationship was reversed. A similar interaction appeared also in the S.D. data, as can be seen in the right part of FIGURE 4.

Reaction Times

Mean individual RTs were calculated for each stress location, each stress type and each level of choice. These means were subjected to a three-way ANOVA, which yielded significant main effects for stress type ($F = 23.49$, d.f. = 1,11, $p < .001$) and number of choices ($F = 62.76$; d.f. = 1,11, $p < .001$), while the interaction between these two factors was nonsignificant. As shown in the left part of FIGURE 5, RT increased as the number of equally likely stress locations increased from 2 to 4, and it was longer under the stress- condition than under the stress+ condition. There was no significant main effect for stress location, but a highly significant stress location by stress type interaction ($F = 8.59$, d.f. = 3,33, $p < .001$). The origin of this interaction clearly appears in the middle part of FIGURE 5, where RTs are shown as a function of

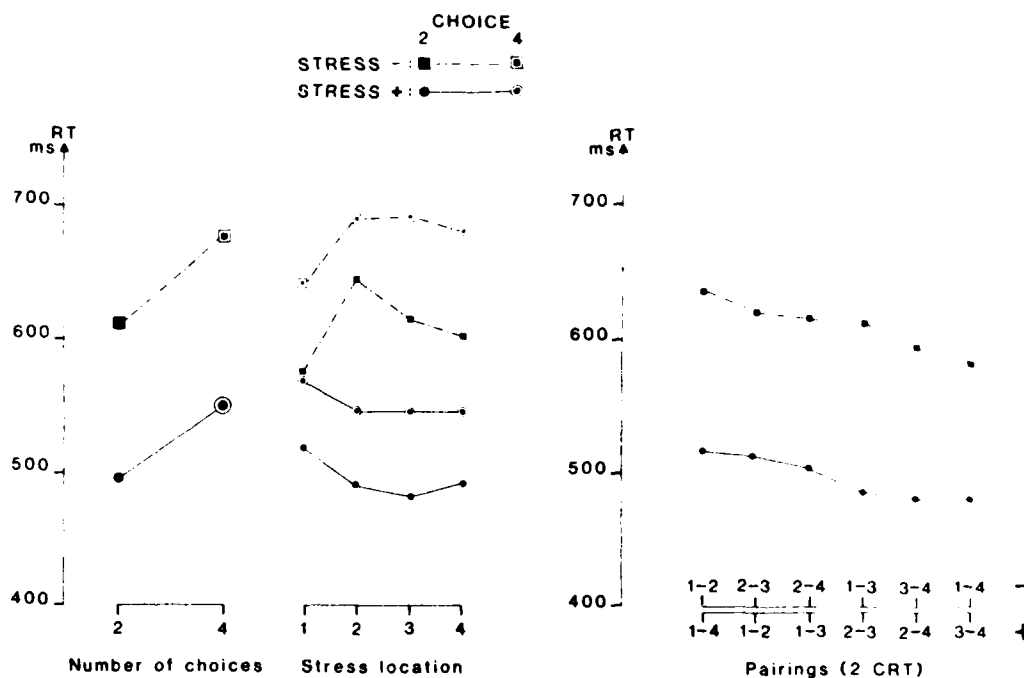


FIGURE 5. (Left) Mean reaction time (RT) as a function of number of choices and stress type. (Middle) Mean RT as a function of stress location, number of choices and stress type. (Right) Mean RT for the pairings in the two-choice condition. Pairings are ranked in decreasing order for each stress type.

stress location. It can be seen that stressing the first tap provoked the longest RT within the stress+ condition, and the shortest RT within the stress- condition. It must be noted, however, that for a given level of choice, RT was always longer for the stress- than for the stress+ condition.

Subsequent partial comparisons showed that stress location was a significant factor within each stress-type condition (for stress+: $F = 3.39$, d.f. = 3,33, $p = .05$; for stress-: $F = 6.36$, d.f. = 3,33, $p < .01$). There was also a significant stress location by number of choices interaction within the stress- condition. This interaction was

mainly due to RT being particularly long under the two-choice condition when the stress was located on the second tap.

RTs obtained under the two-choice condition were also analyzed in order to determine whether pairing of the sequences was by itself a significant source of variance in the two-choice RT. It should be recalled that the six possible pairwise combinations between the four different stress locations were used. Mean individual RTs obtained for the six pairings were subjected to a one-way ANOVA, which showed a significant treatment effect both for the stress+ and stress- condition (the F ratios were, respectively, 3.27 and 2.87, d.f. = 5,55, $p < .05$).

The variations in RT as a function of sequence pairings are shown in the right part of FIGURE 5. RTs were ranked in decreasing order. Only some of these values differed significantly from each other in pairwise *post hoc* comparisons. Under the stress+ condition, the 1-4 pairing differed significantly from the 2-3, 2-4, and 3-4 pairings. Also, the 1-2 pairing differed significantly from the 2-4 pairing. Under the stress- condition, the 1-2 pairing differed significantly from the 1-3, 3-4, and 1-4 pairings. Also, the 2-3 pairing differed significantly from the 3-4 pairing. All other differences were nonsignificant. It is important to note that the ordering of the pairings was clearly different between the stress+ and stress- conditions.

DISCUSSION

This section shall involve two main parts. First, RT data will be evaluated with regard to the advance planning problem. Second, the results concerning the timing of the sequence will be considered.

The Advance Planning Problem

Number of Alternative Stress Locations

As predicted, RT increased as a function of the number of equiprobable locations of the stress within the sequence of four finger taps. Since RT is known to increase when the number of response alternatives increases,^{5,6} a simple interpretation of our results is that stress location differentiates the whole sequence as a particular response. Therefore, under the four-choice condition RT lengthens as compared to the two-choice condition because in the former case the subject must identify the appropriate response within a larger set of response alternatives than in the latter case. Although *sequence identification* does not imply that a detailed program is also established for the *execution* of the sequence, it appears nevertheless as a prerequisite for generating such a program.

Our interpretation assumes that response identification was the main contributing factor to the observed variations in RT, whereas the time needed for stimulus discrimination changed little, if any, as a consequence of the increase in the number of equiprobable *stimuli*. This assumption is based on known evidence that manipulation of the number of equiprobable stimulus-response pairs produces sizable choice RT effects; however, this is true only if the conditions of the experimental procedure make it difficult to identify the responses assigned to the different stimuli.⁷ When response identification is easy and direct, as in numeral-naming tasks, where subjects are simply

asked to read aloud (to name) visually presented digit stimuli, the response latencies do not show much variation as a function of the number of equiprobable digits. When the same digit stimuli are associated each with the movement of a different finger (button press), the movement latencies increase sharply as a function of the number of equally likely stimulus alternatives.^{5,8}

The assumption that in our experiment the manipulation of the number of alternative stress locations primarily affected response identification rather than stimulus discrimination was corroborated by the results of a control experiment in which our subjects performed a numeral-naming task. Although naming latencies turned out to be longer under the four-choice condition than under the two-choice condition, the difference was only one-third of that found when sequences of finger movements were associated with the same digits.⁶

The sequence identification hypothesis states that before the subject starts to execute such rapid movement sequences, he identifies the whole sequence as the required response. An alternative hypothesis would be that, during the RT period, only the first element of the sequence is selected for execution. Under the first-item identification hypothesis RT should depend on the predictability of that first element.

Under the two-choice condition the predictability of the first element of the sequence (that is, whether a high- or a low-force tap) varied as a function of pairings. For the pairings in which the stress was never located on the first tap (2-3, 2-4, 3-4), the first element was fully predictable. For the other pairings (1-2, 1-3, 1-4), the probability that the first tap was a high- or a low-force tap equaled .50. Under the first-item identification hypothesis one would predict sizable RT differences between the pairings in which the first tap was fully predictable and those in which a high- or a low-force tap was equally probable. Moreover, identical RT would be predicted for pairings belonging to the same category. Our data lent no support for either of these predictions.

Under the four-choice condition, the probability that the first tap was a high- or low-force one was .25 and .75 under the stress+ condition, and .75 and .25 under the stress- condition, respectively. The first-item identification hypothesis would predict that sequences starting with a .75 probability tap (under the four-choice condition) would have shorter RT than those in which a low- or high-force tap is equally probable (pairings 1-2, 1-3, 1-4 under the two-choice condition). *Post hoc* analysis of the data lent no support for this prediction; RT under the four-choice condition was longer than under the two-choice condition regardless of the probability of the first tap.

The predictability of the first element of the sequence varied according to pairings and choice conditions, but not according to type of stress. The first-item identification hypothesis would predict identical RT for stress+ and stress- sequences since in either case the subject would have to make the same binary choice between a low- and high-force tap. This hypothesis was not supported by the experimental data.

Since none of the predictions of the first-item identification hypothesis received

⁶In an additional session, subjects were presented with the same digital stimuli used in the main experiment. The task consisted of naming a digit as quickly as possible after its presentation. Simple, two-choice, and four-choice RT conditions were run as described in the METHOD section. Mean naming latencies, together with the RTs for the tapping task, were subjected to an ANOVA. Significant effects were obtained for type of task (naming or tapping) ($F = 28.38$, d.f. = 1,11, $p < .001$) and number of choices ($F = 69.20$, d.f. = 1,11, $p < .001$). The interaction between these two factors was also significant ($F = 11.26$, d.f. = 1,11, $p < .01$). On the average, the four-choice RT was 60.6 msec longer than the two-choice RT for the tapping task, while there was only a 19-msec difference for the naming task.

support from the data, it is safe to conclude that the sequence as a whole, rather than its first element, was identified as the required response during the RT.⁶

Type of Stress

As expected, it took more time for the subject to initiate a sequence in which a tap was weaker than all the others, than to initiate a sequence in which one tap was stronger than all the others. Furthermore, the effect of type of stress and that of the number of alternative stress locations combined additively. From the perspective of Sternberg's additive factor method for analyzing RT,⁹ number of alternative stress locations and type of stress could be regarded as two factors that relate to separate processes in response organization. The first factor may be linked to a sequence identification stage, whereas the second one may be linked to a stage of planning or programming of the response execution.

From the present results, it is difficult to determine the kind of information and the degree of detail contained in the plan for response execution. Timing of the sequence, as revealed by the duration of the intervals between the successive taps, did not depend on type of stress. Recently, Klapp contended that response timing is the main factor that provokes changes in response program complexity and hence in programming time.¹⁰ However, our experimental data indicate that timing was not a critical factor for the lengthening of RT under the stress- condition. The time distribution of motor intensities (forces) over the component elements was very different according to the type of stress. Positive stress involved the generation of a high-force element on the background of a stable low-level force selected for nonstressed taps. It may be that under this condition planning concerned only the force and the time location of the stressed tap, whereas the background low-level force was regulated by more automatic control mechanisms. In contrast, negative stress involved the generation of a low-force element on the background of an unstable high-level force selected for nonstressed taps. Here the planning may have concerned not only the force and time location of the negatively "stressed" element, but also the maintenance of the background force on a unusual high level. This more complicated planning would result then in a longer RT.

Effect of Stress Location

The RT of a sequence beginning with a stressed element was different from that of a sequence beginning with a nonstressed element. Under the stress+ condition, the stressed tap was of high force and the nonstressed taps of low force. The sequences beginning with a high-force element had longer RT than those beginning with a

⁶How is the response sequence identified? It has been suggested⁵ that response identification involves search for the abstract name code of the response. However, such an abstract representation of the response may be of little use for the motor system. If so, the abstract response code must be translated into more concrete representations. We feel that the ultimate identifier of the response is nothing more than the plan or program that permits its execution. It seems likely that in the course of response organization intermediate response representations are generated between the abstract response code and the plan for response execution. The sensory or proprioceptive image of the response could be such a representation.¹² We are presently investigating the hypothesis that the rhythmic response sequences are identified in terms of the acoustic feedback images that are normally associated with the production of such sequences.

low-force element. Under the stress- condition, the stressed tap was of low force and the nonstressed taps of high force. Here again, the sequences beginning with a high-force tap had longer RT than the sequences beginning with a low-force tap. From these results one can conclude that programming a high-force initial element required more time than programming a low-force initial element. This conclusion is in accordance with our previous suggestion that low force was regulated in a more automatic way than high force. However, it must be emphasized again that type of stress effect does not simply derive from the decision concerning the level of force of the initial tap. Sequences that started with a high-force tap under the stress- condition had RTs 100 msec longer than sequences beginning with a high-force tap under the stress+ condition. The same relationship held true for sequences that started with a low-force tap. Thus, type of stress effect would be related to the planning of the sequence as a whole, while the observed stress location effect would be related to the programming of the initial element.

In conclusion, one can suggest that the short rhythmic sequences studied in this experiment are planned in advance of their execution. Planning involves, first, the identification of the sequence as a particular response, and second, the generation of a time-force plan that concerns the sequences as a whole. From the stress location effect we can infer that, in addition to the general time-force plan of the sequence, an executive program is generated that specifies the appropriate force level for the first element of the sequence.

The Timing of the Sequence

In our experiment the subjects had to reproduce a temporal model in which the duration of the intervals was the same. What was actually produced were strings of finger taps in which the intervals showed lawful variations with respect to the model. The structure of these variations could not be apprehended in the movement sequences as they were produced. It became apparent only when the time intervals produced were sorted according to a double classification principle based on position of the interval in the sequence and on position of the interval with respect to stress (or to change in force). This classification of the intervals permitted us to uncover: (1) a time structure linked to interval position in the sequence; and (2) the modulation of that basic structure by the stress.

Interval Modulation by Its Position in the Sequences

The time structure underlying the movement sequences was characterized by the fact that the first and last intervals tended to be equal, while both were considerably longer than the intermediate one. This time structure was force-independent in the sense that it was basically the same whether at the end of the intervals a force increase, a force decrease, or no change in force should occur. In previous experiments the same relationship between intervals was observed in five-tap strings as was demonstrated in the present experiment in four-tap strings. These time configurations are not easily accounted for by models in which an internal clock sends trigger pulses directly to the motor system for producing cadenced finger taps. Although the aim of the present work was not to test any formal model of sequence timing, our results might lend some support for a model in which the role of the internal clock is to provide a temporal reference for movement execution, rather than to trigger each component movement.¹¹

Interval Modulation by Stress

The intervals that immediately preceded and followed an inverted stress were lengthened similarly to those that preceded and followed an ordinary stress. This indicates that the lengthening of the intervals is related to a change in force level, regardless of the direction of that change. Consequently, the effect of the stress on interval duration cannot be explained by the intervention of peripheral (biomechanical) factors responsible for the generation of higher level of force. Instead, the lengthening of the intervals that precede and follow a stressed element seems to be related to central processes that control the appropriate change in force level. This suggests that, in addition to the plan established during the RT, real-time control processes are also involved in the execution of the sequence. These processes could comprise a memory search for the target force level and/or updating the executive program with specification of the new level of force.

SUMMARY

Accentuation involves modulation of motor intensity. It differentiates a movement from others within a motor sequence. Does the serial position of the accent characterize the whole sequence as a particular response? How are the control of time and force coordinated in the motor sequence? Subjects produced sequences of four fingertaps on a key. Time of onset and force of each tap were recorded. Tapping rate was imposed by a string of four clicks delivered at 180-msec intervals before each trial. A flashed digit served as go signal. It indicated to the subject which of the four taps had to be tapped stronger (stress+) or weaker (stress-) than all the others. These conditions were run in separate series. Reaction time (RT) of the sequence increased when the number of equally likely locations of the stress increased from 2 to 4. RT was also longer under the stress- than under the stress+ condition. Tapping intervals were longer before and after the stressed tap than elsewhere in the series. The first and last intervals tended to be longer than the second one. These effects were the same under both stress conditions. The RT data indicate that the motor sequence is identified as a particular response before it starts. Timing is partly force-independent, but is modulated by central processes that control force.

REFERENCES

1. VINCE, M. A. 1948. Corrective movements in a pursuit task. *Q. J. Exp. Psychol.* **1**: 85-103.
2. HENRY, F. M. & D. E. ROGERS. 1960. Increased latency for complicated movements and a "memory drum" theory of neuromotor reaction. *Res. Q.* **31**: 448-458.
3. STERNBERG, S., S. MONSELL, R. L. KNOLL & C. E. WRIGHT. 1978. The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. *In Information Processing in Motor Control and Learning*. G. E. Stelmach, Ed.: 117-152. Academic Press, New York, NY.
4. OSTRY, D. J. 1980. Execution-time movement control. *In* *Tutorials in Motor Behavior*. G. E. Stelmach & J. Requin, Eds.: 457-468. North-Holland, Amsterdam.
5. THEIOS, J. 1975. The components of response latency in simple human information processing tasks. *In* *Attention and Performance*. P. M. A. Rabbitt & S. Dornic, Eds.: 418-440. Academic Press, London and New York.
6. HICK, W. E. 1952. On the rate of gain of information. *Q. J. Exp. Psychol.* **4**: 11-26.

7. BRAINARD, R. W., T. S. IRBY, P. M. FITTS & E. A. ALLUISE. 1962. Some variables influencing the rate of gain of information. *J. Exp. Psychol.* **63**: 105-110.
8. THEIOS, J. 1973. Reaction time measurement in the study of memory processes: Theory and data. In *The Psychology of Learning and Motivation: Advances in Research and Theory*, G. H. Bower, Ed. Vol. 7: 43-85. Academic Press, New York and London.
9. STERNBERG, S. 1969. The discovery of processing stages: Extension of Donder's method (Attention and Performance: II). *Acta Psychol.* **30**: 276-315.
10. KIAPP, S. T. 1977. Reaction time analyses of programmed control. *Exercise Sport Sci. Rev.* **5**: 231-253.
11. SHAFFER, L. H. 1981. Performances of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychol.* **13**: 326-376.
12. GREENWALD, A. G. 1970. Sensory feedback mechanisms in performance control with special reference to the ideomotor mechanisms. *Psychol. Rev.* **77**: 73-99.

7. BRAINARD, R. W., T. S. IRBY, P. M. FITTS & E. A. ALLISEL. 1962. Some variables influencing the rate of gain of information. *J. Exp. Psychol.* **63**: 105-110.
8. THEIOS, J. 1973. Reaction time measurement in the study of memory processes: Theory and data. *In* *The Psychology of Learning and Motivation: Advances in Research and Theory*. G. H. Bower, Ed. Vol. 7: 43-85. Academic Press, New York and London.
9. STERNBERG, S. 1969. The discovery of processing stages: Extension of Donder's method (Attention and Performance: II). *Acta Psychol.* **30**: 276-315.
10. KLAPP, S. T. 1977. Reaction time analyses of programmed control. *Exercise Sport Sci. Rev.* **5**: 231-253.
11. SHAFFER, L. H. 1981. Performances of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychol.* **13**: 326-376.
12. GREENWALD, A. G. 1970. Sensory feedback mechanisms in performance control with special reference to the ideo-motor mechanisms. *Psychol. Rev.* **77**: 73-99.

Motor Disorder and the Timing of Repetitive Movements^a

ALAN M. WING

*Medical Research Council
Applied Psychology Unit
Cambridge CB2 2EF, England*

STEVEN KEELE AND DAVID I. MARGOLIN

*Cognitive Neuropsychology Laboratory
Good Samaritan Hospital and Medical Center
Portland, Oregon 97210*

INTRODUCTION

In psychology, there are a number of instances in which accounts of human information processing have been advanced by the study of pathologic states. In the case of language, for example, the study of aphasia resulting from brain disease not only provided the first links toward understanding language and brain anatomy, but has also contributed to the psychological understanding of language in terms of the functional mechanisms involved. This is true because language breakdown is often selective; it is quite common to observe patients with some aspects of language function severely disrupted, while other aspects remain relatively intact. An example is provided by the contrasting selective deficits of Broca's and Wernicke's aphasia with damage to anterior and posterior areas of the brain.

In the present paper, we consider a theoretical model of the timing of repetitive movements that was developed by Wing and Kristofferson¹ as an account of data from normal subjects. The model postulates the existence of two separate processes contributing to the variability observed in the intervals between a stream of supposedly regular responses. We ask: In motor disorders of the central nervous system can a selective deficit be demonstrated in one or other of these processes? We address this question by presenting a detailed study of finger tapping by a patient with Parkinson's disease.

Parkinson's disease is a degenerative disease affecting the dopaminergic pathways of the basal ganglia. Symptoms of the disease affecting arm movements can include rest tremor (a tremor seen when the arm is held in a given posture, of larger amplitude and lower frequency than normal physiological tremor), rigidity (heightened resistance of the arm to passive manipulation), and bradykinesia (slowness of voluntary movement). The basal ganglia are bilateral structures of the midbrain involved in control of movement in the contralateral side of the body, but cases of predominantly unilateral symptoms have been reported.²⁻⁴ We describe a case of hemi-parkinsonism in which a primary symptom was bradykinesia of the dominant (right) hand, which

^aThis research was supported in part by Grant NSF BNS 8119274 from the National Science Foundation (to S.K.), by Individual Research Award F32 NS0 6788 from the National Institute for Neurological and Communicative Disorders and Stroke (to D.I.M.), and by a grant from the American Parkinson Research Foundation.

AD-A150 031

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES VOLUME 423
TIMING AND TIME PER. (U) NEW YORK ACADEMY OF SCIENCES
NY J GIBBON ET AL. 13 MAY 84 N00014-84-G-0130

3/3

UNCLASSIFIED

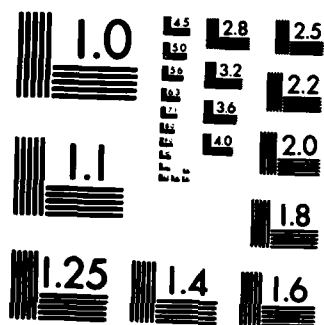
F/G 5/10

NL

END

FILED

ETC



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

had deleterious effects on the timing of movements in that hand. In this paper we compare the application of the Wing and Kristofferson model to timing data from each hand.

Consider a task in which the subject is instructed to produce a series of responses, such as tapping the index finger of one hand, in synchrony with a periodic auditory stimulus. After a few taps to get into phase with the pacing stimulus (an account of the processes involved may be found in Michon⁵), the auditory pacer is discontinued and the subject is asked to continue to tap at the same rate. Under these conditions subjects are able to produce long series of "free" responses with stable interresponse intervals (IRIs) with little drift in mean.¹

Wing and Kristofferson's model of self-paced responding¹ (FIG. 1) proposes that departures from periodic responding may arise from imprecision in a hypothetical timekeeper and from temporal "noise" in the execution of responses triggered by the timekeeper. At time intervals, C , the timekeeper emits pulses, each of which initiates a motor response. A motor output delay, D , intervenes between the initiation and occurrence of the overt response, R . Each interresponse interval is thus the sum of a timekeeper interval plus the difference in motor delays associated with the initiating and terminating responses:

$$I_j = C_j + D_j - D_{j-1}. \quad (1)$$

Provided timekeeper intervals and response delays are both sequences of independent random variables and are mutually independent, it can be shown⁶ that the IRI autocovariance function may be expressed as:

$$\text{autocov}_I(0) = \text{var}(I) = \text{var}(C) + 2 \text{var}(D) \quad (2)$$

$$\text{autocov}_I(1) = -\text{var}(D) \quad (3)$$

$$\text{autocov}_I(k) = 0, \quad k > 1. \quad (4)$$

Thus, the model predicts negative covariation between immediately successive IRIs ($\text{autocov}_I(1) < 0$). Chance variation in any particular delay, D , about the mean motor output delay will tend to produce deviations of opposite signs in I_j and I_{j+1} about the mean IRI. The covariance between adjacent intervals, involving the products of such deviations, thus tends to be negative. Intervals that are separated by at least one intervening IRI and so do not share a common boundary should have zero covariance.

In previous experimental work,⁶ various lines of research have provided evidence in support of the two-process model. First, estimates of the covariance between adjacent IRIs are usually less than zero.¹ Second, Wing⁷ found that changes in the mean interval between responses only affect timekeeper variance (which may be estimated by solving for $\text{var}[C]$ in Equations 2 and 3). Third, Wing^{8,9} has found that, with changes of effector, timekeeper variance is more stable than motor delay variance (however, another parameter was added to the model, a point we will return to in the DISCUSSION).

If timing of the left and right hands of a patient with symptoms of parkinsonism affecting one side more than the other are compared, the theoretically interesting question is whether left-right differences in the variance estimates of the hypothesized underlying processes provide evidence that the processes are dissociated. If there is greater timing variability in the tapping of one hand compared with that of the other, will this be attributable to elevated motor-delay variance (larger, more negative lag-one variance) or to elevated timekeeper variance (larger IRI variance with no difference in lag-one covariance)? Or, will there be changes in both?

The average velocity of the left hand was greater than that of the right. In the left hand the velocity showed a slight increase from the first (1.15 m/sec) to the tenth letter, (1.36 m/sec). In contrast there was a decrease in velocity in the right hand from the first (0.66 m/sec) to the tenth letter (0.42 m/sec).

2. *Reciprocal aimed tapping.* The patient was asked to use a pencil in the right or the left hand to tap alternately in pairs of targets 5 mm or 20 mm wide, 51 mm or 205 mm apart. The number of taps made within a 10-sec trial was recorded for each of the four combinations of width and distance. Over eight trials for each combination the time per tap for the left hand was reliably less than that for the right hand (TABLE 1). There was a reliable hand-by-width interaction; the right hand failed to speed up movements as much as the left hand to capitalize on the easiness of the wide targets.

3. *Purdue pegboard.* A task requiring perceptual-motor coordination similar to the previous task is the Purdue pegboard (Science Research Associates). Small cylindrical pegs must be picked up and placed in a row of drilled holes. From the published norms for this test, it was established that M.F.'s performance was below the first percentile for the right hand. But even with her left hand, her score was only at the fifth percentile. Although M.F. is used as her own control in the experiment described below, it is important to recognize that left-hand performance is not unimpaired. It might be noted that studies have also indicated that changes associated with Parkinson's disease tend to be bilateral even if markedly asymmetric.³

4. *Reaction time.* In a task developed for studying possible effects of line-drawing movement complexity on reaction time, it was found that M.F.'s simple and choice reaction were slower than normal controls. The median latency of the right hand was some 16 msec longer than that of the left hand.

5. *Maximum rate of tapping.* The maximum tapping rate was slower in the right hand (315 msec per tap) than in the left hand (215 msec per tap). Both values are longer than the normal range.¹¹

EXPERIMENT

Task

The subject was seated with the arm used for tapping resting on a table. A touch plate was placed in front of the subject who placed her hand on it, palm down. Tapping the index finger on the touch plate completed a low-voltage electrical circuit, which provided a pulse to an Apple II computer, allowing time of contact to be measured to the nearest millisecond.

On each trial a series of 20 brief, clearly audible tones were presented at regular intervals under computer control. The standard intervals used were 450 and 550 msec. The subject was instructed to use movements at the carpo-metacarpophalangeal joint to tap in synchrony with these tones and to continue tapping at the same rate when the

TABLE 1. Reciprocal Tapping Task: Time per Tap (msec) as a Function of Target Distance and Width for Patient M.F.

	Left Hand		Right Hand	
	Wide	Narrow	Wide	Narrow
Near	328	414	416	451
Far	439	549	545	623

END

FILMED

3-85

DTIC