Volume V

QUANTITATIVE ANALYSES OF BEHAVIOR

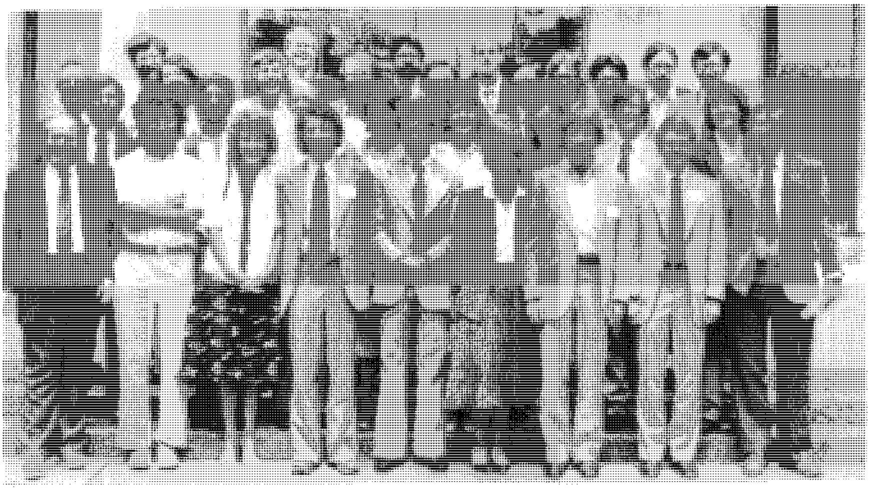
The Effect of Delay and of Intervening Events on Reinforcement Value

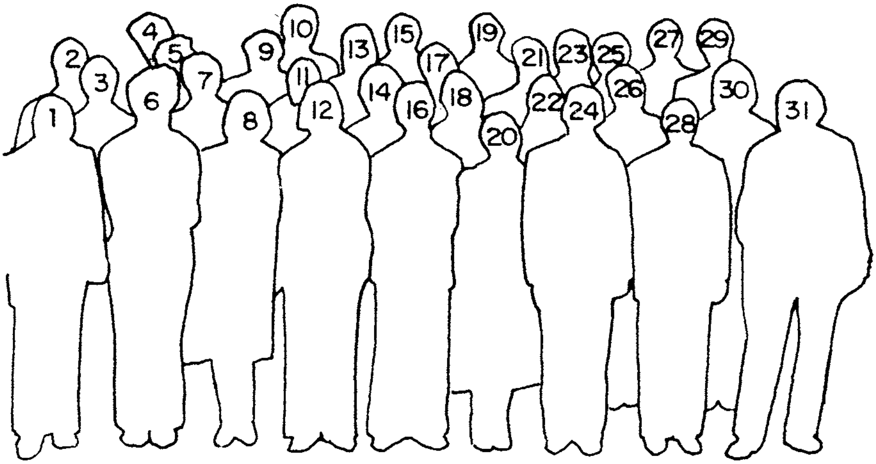
Edited by

MICHAEL L. COMMONS
JAMES E. MAZUR
JOHN A. NEVIN
HOWARD RACHLIN

Psychology Press
Taylor & Francis Group

QUANTITATIVE ANALYSES
OF BEHAVIOR





1. Michael L. Commons 2. John A. Nevin 3. Robert Epstein 4. William Baum 5. Unknown 6. Seth Roberts 7. Michael C. Davison 8. Dianne C. McCarthy 9. Howard Rachlin 10. Allan R. Wagner 11. George Ainslie 12. Edmund J. Fantino 13. John Gibbon 14. Nureya Abarca 15. Jack J McDowell 16. Telmo E. Peña-Correal 17. Richard J. Herrnstein 18. Alexandra W. Logue 19. Jay Moore 20. Monica L. Rodriguez 21. Robert Eisenberger 22. Debra J. Spear 23. Michael Woodford 24. Philip N. Hineline 25. Mark S. Snyderman 26. Richard L. Shull 27. Leon R. Dreyfus 28. Kennon A. Lattal 29. James E. Mazur 30. J. Gregor Fetterman 31. D. Alan Stubbs

QUANTITATIVE ANALYSES OF BEHAVIOR

The Effect of Delay and of Intervening Events on Reinforcement Value

Volume V

Edited by

MICHAEL L. COMMONS
Harvard University

JAMES E. MAZUR
Harvard University

JOHN A. NEVIN
University of New Hampshire

HOWARD RACHLIN
State University of New York
Stony Brook

Psychology Press
Taylor & Francis Group

New York London

First Published by
Lawrence Erlbaum Associates, Inc., Publishers
365 Broadway
Hillsdale, New Jersey 07642

Transferred to Digital Printing 2009 by Psychology Press
270 Madison Ave, New York NY 10016
27 Church Road, Hove, East Sussex, BN3 2FA

Library of Congress Cataloging in Publication Data
Library of Congress Catalog Card Number: 81-2654
ISBN 0-89859-800-1

Volumes in the **QUANTITATIVE ANALYSES OF BEHAVIOR** series:

- Volume I: DISCRIMINATIVE PROPERTIES OF REINFORCEMENT SCHEDULES. Edited by Michael L. Commons, *Harvard University* and John A. Nevin, *University of New Hampshire*
- Volume II: MATCHING AND MAXIMIZING ACCOUNTS. Edited by Michael L. Commons, *Harvard University,* Richard J. Herrnstein, *Harvard University,* and Howard Rachlin, *State University of New York, Stony Brook*
- Volume III: ACQUISITION. Edited by Michael L. Commons, *Harvard University,* Richard J. Herrnstein, *Harvard University,* and Allan R. Wagner, *Yale University*
- Volume IV: DISCRIMINATION PROCESSES. Edited by Michael L. Commons, *Harvard University,* Richard J. Herrnstein, *Harvard University,* and Allan R. Wagner, *Yale University*
- Volume V: THE EFFECT OF DELAY AND OF INTERVENING EVENTS ON REINFORCEMENT VALUE. Edited by Michael L. Commons, *Harvard University,* James E. Mazur, *Harvard University,* John A. Nevin, *University of New Hampshire,* and Howard Rachlin, *State University of New York, Stony Brook*
- Volume VI: FORAGING (in press). Edited by Michael L. Commons, *Harvard University,* Alejandro Kacelnik, *Oxford University,* and Sara J. Shettleworth, *University of Toronto*

All volumes available from LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS

CONTENTS

PART II *SINGLE SCHEDULES*

PART III *CONCURRENT SCHEDULES AND CONCURRENT-CHAIN SCHEDULES*

ABOUT THE EDITORS

Michael Lamport Commons is a research associate in the Department of Psychology and Social Relations at Harvard University. He did his undergraduate work at the University of California at Berkeley, and then at Los Angeles, where in 1965 he obtained a B.A. in mathematics and psychology. In 1967 he received his M.A. and in 1973 his Ph.D. in psychology from Columbia University. Before coming to Harvard University in 1977, he was an assistant professor at Northern Michigan University. He has co-edited *Quantitative Analyses of Behavior: Volume I, Discriminative Properties of Reinforcement Schedules, Volume II, Matching and Maximizing Accounts, Volume III, Acquisition, Volume IV, Discriminative Processes,* and *Beyond Formal Operations: Late-Adolescent and Adult Cognitive Development.* His area of research interest is the quantative analysis of the construction and the understanding of reality, especially as they affect decision processes. This includes the subarea of utility, along with such topics as perception and knowledge of causal relations and of value, and how it develops in infrahumans and humans across the lifespan.

James E. Mazur received his B.A. from Dartmouth College in 1973 and his Ph.D. from Harvard University in 1977. He has served on the editorial board of the *Journal of the Experimental Analysis of Behavior,* and is currently a consulting editor of the *Journal of Experimental Psychology: Animal Behavior Processes.* He has contributed several articles to these and other journals, and is the author of a recently published textbook, *Learning and Behavior.* He is now an associate professor of psychology at Harvard University, where he has taught since 1980. His current research concerns how variables, such as delay, amount, and probability of reinforcement, influence choice.

John Anthony Nevin received his B.E. in mechanical engineering from Yale University in 1954 and his Ph.D. in psychology from Columbia University in 1963. He taught at Swarthmore College from 1963 until 1968, and at Columbia University from 1968 until 1972. He served as an editor of the *Journal of the Experimental Analysis of Behavior* from 1979 to 1983 and co-editor of *Quantitative Analyses of Behavior: Volume I, Discriminative Properties of Reinforcement Schedules.* His current research interests include schedules of reinforcement, stimulus control, conditioned reinforcement, behavioral momentum, animal psychophysics, and nuclear disarmament.

Howard Rachlin received his B.E. in mechanical engineering from Cooper Union in 1956, his M.A. in psychology from the New School for Social Research in 1962, and his Ph.D. in psychology from Harvard University in 1965. He taught at Harvard from 1965 until 1969. He is now a professor of psychology at the State University of New York at Stony Brook, where he has been teaching since 1969. His interests include self-control, the interaction of economics and psychology, the matching law, the psychology of eating and drinking, and the philosophy of psychology.

LIST OF CONTRIBUTORS

Nureya Abarca
Department of Psychology
University of California at San Diego
Now at:
Escuela de Psicología
Pontificia Universidad Católica de Chile

George Ainslie
VA Medical Center
and
Jefferson Medical College

Michael C. Davison
Department of Psychology
The University of Auckland

Leon R. Dreyfus
Department of Psychology
University of Maine
Now at:
Department of Psychology
Loyola University

Roger M. Dunn
Department of Psychology
University of California at San Diego
Now at:
Department of Psychology
San Diego State University
Imperial Valley Campus

Robert Eisenberger
Department of Psychology
University of Delaware

Edmund J. Fantino
Department of Psychology
University of California at San Diego

John Gregor Fetterman
Department of Psychology
University of Maine
Now at:
Department of Psychology
Norwich University

Philip N. Hineline
Department of Psychology
Temple University

Kennon A. Lattal
Department of Psychology
West Virginia University

Alexandra W. Logue
Department of Psychology
State University of New York at Stony Brook

Fred A. Masterson
Department of Psychology
University of Delaware

Benjamin C. Mauro
Department of Psychology
State University of New York at Stony Brook
Now at:
Department of Psychology
Temple University

James E. Mazur
Department of Psychology and Social Relations
Harvard University

Dianne C. McCarthy
Department of Psychology
The University of Auckland

Jack J McDowell
Department of Psychology
Emory University

Telmo E. Peña-Correal
Department of Psychology
State University of New York at Stony Brook
Now at:
Departamiento de Psicología
Universidad de los Andes

Monica L. Rodriguez
Department of Psychology
State University of New York at Stony Brook

Richard L. Shull
Department of Psychology
The University of North Carolina

Mark Snyderman
Department of Psychology and Social Relations
Harvard University

Debra J. Spear
Department of Psychology
The University of North Carolina

Frank J. Sodetz
Walter Reed Army Institute of Research
Now at:
United States Army Medical Component
Armed Forces Research Institute of Medical Sciences

D. Alan Stubbs
Department of Psychology
University of Maine

William Vaughan, Jr.
Department of Psychology and Social Relations
Harvard University

K. Geoffrey White
Department of Psychology
Victoria University
Now at:
Department of Psychology
The University of Otago

PREFACE

The study of behavior has consisted of a number of somewhat separate traditions. One tradition, starting with Thorndike and then continuing with Skinner, has analyzed experimentally the control of behavior by events that occur subsequent to it. A second tradition, starting with Bechterev and Pavlov and coming down to the present through Watson, Hull, Spence, and others, has analyzed the control and transfer of control by events that precede behavior.

After the 1920s both approaches became more quantitative. In the experimental analysis of behavior, quantifiable variables, such as the rate of responding, were used to represent the behavioral outcomes. At the same time, more elaborate quantitative studies were carried out in the Hullian approach. Quantifiable measures, such as response probability and latency, were introduced. In that period, and extending through the 1950s, mathematical models were developed by Hull, Spence, Estes, Bush and Mosteller, and Logan, among others. Both groups carried out some parametric studies in the tradition of psychophysics. By the early 1960s mathematical psychology had developed to the point where it could deal with problems from a number of domains. In each domain, explicit mathematical models were proposed for the processes by which performances were acquired and maintained within that domain. Although the models generated a number of experiments, they were of limited generality.

''Quantitative analysis'' now generally refers to the fact that theoretical issues are represented by quantitative models. An analysis is not a matter of fitting arbitrary functions to data points. Rather, each parameter and variable in a set of equations represents part of a process that has both a theoretical and an empirical interpretation. Quantitative analysis has forced researchers to represent explicitly their notions and to be economical in the number of parameters that must be estimated. The matching law, a model of maintained performance, is one example from the analysis-of-behavior tradition. The Rescorla–Wagner model of acquisition processes is a second example. These models represent effects of interactions of environmental and behavioral events. Because neither model requires otherwise, the possibility exists that both the organism and the environment modify each other. The rules of such interaction may be represented by an arithmetic that accounts for the results from a large class of studies.

The models are designed to account for the maximal amount of variance found in a number of experimental situations to which the processes described by a given one of those models apply. Some parameter estimates should be the same regardless of the situation. The adequacy of a model can be tested by examining how well that model fits the data or by comparing obtained data to the theoretically simulated values. These methods are to be contrasted with the testing of relatively simple hypotheses. Because the models can be quite complex, however, only portions of them are tested by single sets of studies. As in other areas of science, looking for the generality of a formulation has made these models more testable. Independent routes of verification are possible because of the increased scope of the models.

The volumes in the present series have been written for behavioral scientists. Those concerned with issues in the study of how behavior is acquired and then allocated in various environments—biologists, psychologists, economists, anthropologists, and other researchers, as well as graduate students and advanced undergraduates in those areas—should find volumes in this series to be state-of-the-art readers and reference works. They are also intended for use in seminars.

Each volume of the series examines a particular topic that has been discussed at the annual Symposium on Quantitative Analyses of Behavior held at Harvard University. The topic of Volume I was the discrimination of schedules of reinforcement. It was chosen because it represents an area that has been highly quantified through the application of psychophysical methods and analyses of maintained performances.

Volume II explored matching and maximizing accounts of the allocation of behavior, another area that has been highly quantified. It explored the generality of such formulations and how they apply to animal behavior in both the field and the laboratory and to human behavior in choice situations in economics.

Acquisition models and data were considered in Volume III. These models dealt with the roles that various events play in different conditioning situations and how those events interact to produce conditioning or to retard it. Aspects of the conditioning situation were considered that go beyond the simple notions of temporal contiguity between the stimulus to be conditioned and the unconditioned stimulus.

Volume IV presented studies of discrimination processes. How discriminations are acquired and the role of various events within the discrimination situation were considered.

This volume, V, addresses the topic of how reinforcement value is affected by delay and intervening events. Self-control studies are also presented and discussed.

Volume VI will address issues in foraging. Included will be an examination of optimal-foraging theory and its alternatives, as well as an examination of how the detectibility of prey controls the choice to pursue those prey.

Volume VII will address the biological determinants of reinforcement and memory.

Tentative future volumes will include Volume VIII, the topics of which are pattern recognition and concepts in animals, people, and machines, Volume IX, whose topic will be economic approaches to human and animal choice, and Volume X, which will deal with stimulus control.

The contents of the present volume were first prepared for and presented at the Fifth Symposium on Quantitative Analyses of Behavior, held at Harvard University on June 6 and 7, 1982. Subsequently, a portion of the chapters has been revised, updated, and rearranged into the four topical parts found herein.

The symposium out of which this fifth volume has arisen was organized by Michael L. Commons, Richard J. Herrnstein, James E. Mazur, John A. Nevin, and Howard Rachlin. The symposium was supported in part by the Society for the Quantitative Analyses of Behavior, The Department of Psychology and Social Relations at Harvard University, and by the Dare Association, Inc.

In 1982 local arrangements were made by Patrice M. Miller and Dean Gallant, with assistance from Theodore L. Allen, Michael Armstrong-Roche, Brian D. Cabral, Mark Constas, Martin N. Davidson, Patricia S. French, Wilson Fong, José Gabilondo, and Benjamin Singer.

For help in editing the chapters we thank Charlotte Mandell and William Vaughan, Jr. For help in reviewing the chapters for stylistic and organizational improvements we thank the staff of the Dare Institute.

Michael L. Commons
Harvard University

I

STUDIES OF DELAY AND CHOICE: MOLAR AND MOLECULAR CONCEPTIONS OF CHOICE

1

Duration Comparison and the Perception of Time

J. Gregor Fetterman
Leon R. Dreyfus
University of Maine at Orono

The study of time perception has a long and varied history in psychology. Research on human time perception has used a diverse group of procedures and independent variables including the effects of different stimuli on time judgments, the psychophysics of time, the development of time perception, the role of cognitive and emotional factors on the experience of time, judgments of "normal" and "abnormal" subjects, the effects of drugs, judgments of simultaniety and succession, and the perception of rhythm (see for example reviews by Boring, 1942; Doob, 1971; Fraisse, 1963, 1978, 1984; Frankenhauser, 1959; Ornstein, 1969; Woodrow, 1951; and the bibliography by Eisler, Linde, Throeng, Lazar, Eisler, & Hellstrom, 1980). Although less well studied, research on animals' time perception has almost as long a history as the human time perception work. Discussions of the role of temporal variables in learning go back at least to Pavlov (1927), with references to the role of temporal factors appearing periodically in some of the older animal learning literature (e.g., Anderson, 1932; Cowles & Finan, 1941; Sams & Tolman, 1935; Skinner, 1938).

Interest in time perception in animals stems in part from the role that temporal discriminations are presumed to play in learning situations, where temporal regularities between behavior and environmental events may occur. For example, behavior under temporally defined reinforcement schedules (e.g., fixed-interval schedules, free-operant avoidance schedules, and differential-reinforcement-of-low-rate (DRL) schedules) suggests the possibility that temporal discriminations may be formed and may contribute to performance (see for example Church, 1978; Gibbon, 1977; Richelle & Lejune, 1980; Staddon, 1974). Experiments on time perception with animals also permit comparisons of different species with different evolutionary and experiential histories. The animal studies complement research on humans, a species with a history of counting and using time pieces.

Research on animals' time perception falls under two general categories (Stubbs, 1979): procedures related to the time-based schedules of reinforcement such as DRL schedules, fixed-interval schedules, and the like (e.g., Catania, 1970; Platt, 1979), and psychophysical trials procedures like those used with humans (e.g., Church & Deluty, 1977; Stubbs, 1968, 1979). With the DRL schedule, for example, a response is reinforced only if it is delayed from the previous response by a minimum time; the behavior that results from this temporal restriction on responding suggests a temporal discrimination. With the psychophysical trials procedures, animals may be trained to make one response if the duration of a stimulus is short (2 sec for example), but to make a second response if the duration is long (for example, 6 sec). Accuracy is generally a function of the relative difference between the two stimulus durations (e.g., Stubbs, 1968).

Although a diverse group of procedures have been used, all the methods used to study time perception in animals share a common feature. This feature distinguishes the animal methodology from the more complex procedures used to study human time perception. In all these procedures in the animal timing literature, the consequences for responding depend on a fixed temporal criterion or cutoff point. The onset of the interval is marked by a specific event, which might be the onset of a stimulus, or some aspect of the animal's behavior. With fixed-interval schedules the interval typically begins with the end of the prior reinforcer; with DRL schedules, the interval is begun with each response. In both cases, however, a response is reinforced only when a specific interval of time has elapsed. Choice procedures often involve different durations to be judged, but in these situations there is one cutoff or criterion interval such that shorter durations occasion one set of consequences and longer durations another.

In contrast to the previous experiments, the present experiments used a discrimination task that differed from the earlier work in two basic ways: (a) The task involved the presentation of two stimulus durations rather than one, and (b) the consequences for responding depended on the *difference* between the two durations rather than on the difference between a duration and a single fixed-criterion interval. Specifically, pigeons were presented with two key-light durations in succession, and then two choice keys were lit. One response was reinforced if the first duration was shorter than the second, and the alternate response was reinforced if the first duration was longer than the second.

The experiments bear on issues related to time perception. The paired-comparison task parallels one that is commonly used to study time perception with humans (e.g., Doob, 1971), and thus it permits comparisons between human and animal data. In addition, the experiments have implications for theories of animal time perception, such as the internal clock model of Church and his associates (e.g., Church, 1978; Meck & Church, 1983; Roberts & Holder, 1982). The experiments were not designed as a test of any model, but they do bear on models of timing.

The experiments also have implications for memory and discrimination learning in animals. Memory could be a factor because the stimulus in any duration task lasts over time. The paired-comparison task is of particular interest in this regard because choice depends on two successively presented durations, with the first always delayed from choice. Finally, the experiments have implications for discrimination learning because the paired-comparison task requires a relational judgment. We return to these issues after the data have been presented.

GENERAL METHOD

Figure 1.1 diagrams the general procedure used in both experiments. A trials procedure was used. At the beginning of each trial the center key of a three-key

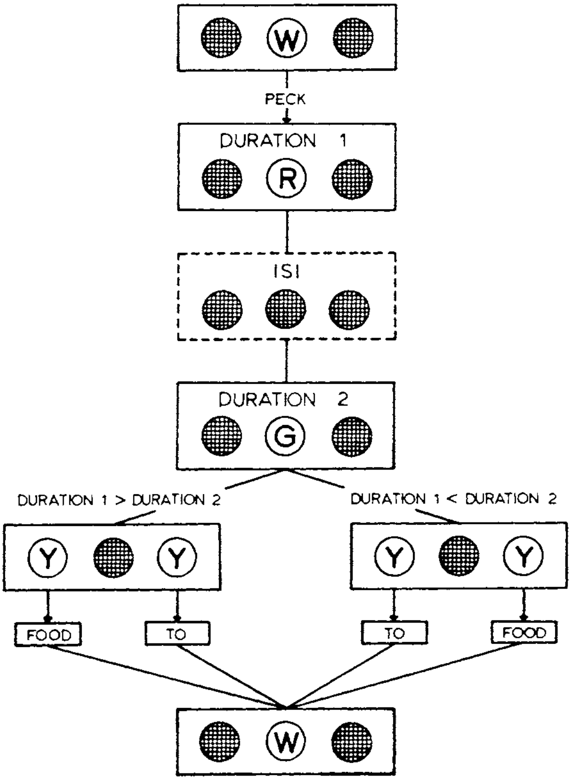


FIG. 1.1. A schematic of the procedure. Each circle represents a response key. The labels W, R, G, and Y stand for key colors white, red, green, and yellow. (From Fetterman & Dreyfus, 1986 by permission of Elsevier Biomedical Press B.V.)

pigeon chamber was lit by white light. A peck to this key changed the color to red and initiated the first duration. The light remained on for a given duration, which changed from trial to trial, and then went off independently of behavior. Under most conditions the offset of red was followed immediately by the onset of green, which demarcated the second duration. Under some conditions, an interstimulus interval, during which the key lights were off, separated the two durations. In either case the green stimulus remained on for a specified duration, changing from trial to trial, and then went off independently of behavior. Offset of green was followed by the onset of two side key lights. A response to one side key was reinforced if the first duration was shorter than the second, whereas the alternate response was reinforced if the first duration was longer than the second. Correct choices were reinforced according to a random-ratio two schedule. Every other response, on the average, produced 3-seconds access to food, followed by a 12-second intertrial interval. Correct responses that did not result in access to food simply produced a 15-second intertrial interval, as did all incorrect choices. The key lights and a houselight that was normally on during trials were all off during food delivery and intertrial intervals. Sessions were conducted 6 days a week with each session lasting until 80 reinforcers had been delivered.

EXPERIMENT 1

For the first experiment,[1] two groups of four durations were used to construct the duration pairs. The first group of durations included durations of 0.5, 1, 2, and 4 seconds. Each of these times served as the first (red) and second (green) duration of a pair in such a way that each duration was combined with the other three in all possible ways: 0.5 sec of red was followed by 1, 2, or 4 seconds of green; 1 second was followed by 0.5, 2, or 4 seconds, etc. The different pairs were arranged in an irregular order, with each pair occurring equally often. Once performance became stable under these conditions (approximately 50 sessions exposure), a new group of durations was used for the duration pairs: 2, 4, 8, and 16 seconds. These durations were presented in the 12 different combinations of each duration followed by the other 3, as was the case under the first condition.

Generalization tests were given at the end of training under each range of durations. Novel duration pairs were introduced as probes that were intermixed with the standard training durations. Responses to these novel pairs were never reinforced with food. Instead choices simply initiated the intertrial interval. The novel pairs consisted of cases where the two durations were the same (e.g., 2 seconds followed by 2 seconds), consisted of durations that lay within the range of training durations (e.g., 3 seconds followed by 2 seconds), and consisted of durations that lay outside the range (e.g., 20 seconds followed by 16 seconds).

---

[1] Portions of Experiment 1 were reported at the Eastern Psychological Association, Baltimore, 1982.

Normally the red stimulus was followed immediately by green, so a final set of conditions imposed an interstimulus interval between red and green in order to separate the first duration from the second. The interstimulus intervals were 0, 2, and 5 seconds under one condition and 0, 10, and 30 seconds under a second condition. Interstimulus intervals were imposed only under the second range of durations.

Figure 1.2 provides the basic information on discrimination performance with both sets of duration pairs. The left side shows performance when the durations were 0.5, 1, 2, and 4 seconds. The right side shows performance for durations of

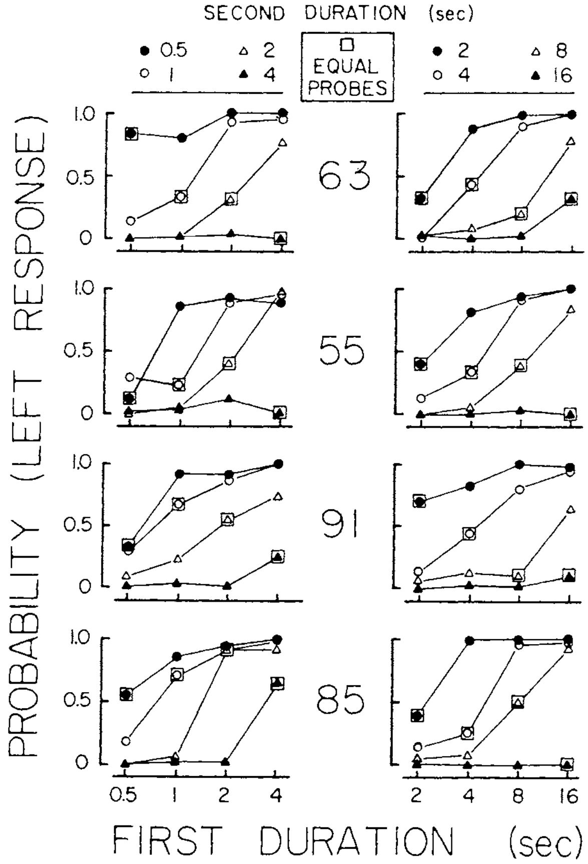


FIG. 1.2. Probability of a left-key response (reporting the first duration to be longer than the second) as a function of the first duration of a pair. Data are presented when two different ranges of durations were used. Data for unequal pairs were computed from totals of the last three sessions of training. Data for equal probe pairs were taken from the first day on which these test pairs were introduced. (From Fetterman & Dreyfus, 1986 by permission of Elsevier Biomedical Press B.V.)

2, 4, 8, and 16 seconds. Data are included for the probe trials in which the first and second durations were equal. Figure 1.2 gives the probability of a left response (duration 1 longer than duration 2) as a function of the value of the first duration of a pair at each second duration. Different values of first durations are ordered along the horizontal axis whereas different values of second duration are represented by different symbols. The filled circles, for example, show performance when 0.5, 1, 2, and 4 second durations were followed by 0.5 seconds. The figure shows that the animals generally were correct in reporting that the first duration was longer. The data in this figure indicate that performance was generally accurate, with the functions showing an abrupt transition from a low to high probability as the first duration changed from shorter to longer than the second. Accuracy increased as the difference between the two durations increased; as the duration difference increased the probabilities approached 0, when the first duration was shorter than the second, or 1.0 when the first was longer. Performance was similar across both ranges of durations and accuracy was high under both conditions (approximately 85% for all pairs of durations).

When equal durations were arranged on probe trials (symbols surrounded by squares), probability measures were variable, ranging around 0.50. This result is not unexpected because neither duration was longer. Although performance was variable, roughly two thirds of the points (21 of 32 cases) fall below 0.50, meaning that the pigeons more often than not reported the first of two equal durations as being shorter. The finding, which could be viewed as a negative time-order error, is sometimes obtained with human time judgments (e.g., Allan, 1977).

Novel duration pairs were introduced under both duration ranges. Figure 1.3 shows accuracy measures for these transfer tests following training on the first group of durations (top) and on the second group (bottom). The figure shows performance averaged for the four pigeons; the vertical bars indicate plus and minus one standard error of the mean. Figure 1.3 shows that transfer to novel duration pairs was generally quite good; accuracy measures above 70% were obtained for individual pigeons in nearly three-quarters of the cases. In some instances, accuracy was related to the relative difference between pair members. Accuracy was high with relatively large differences (for instance, 6 vs. 2 seconds and 12 vs. 6 seconds) but was generally lower when the difference was small (for example, 4 vs. 3 seconds and 8 vs. 6 seconds).

Accuracy was at or below chance level on some probe pairs (e.g., 4 vs. 3 seconds and 20 vs. 16 seconds), and these results implicate factors apart from the relative temporal difference of the two durations. Two factors may account for the relatively poor transfer to some of the novel duration pairs. First, accuracy was low only on problems where the first duration was longer than the second, an outcome consistent with the negative time-order errors observed with equal duration pairs. The poor performances may have resulted from a differential weighting of the values of the two durations, with the nominal value of the first

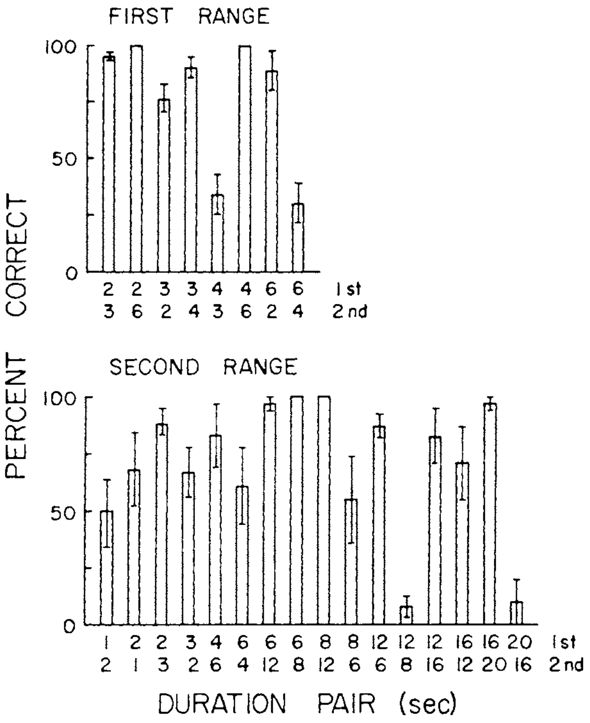


FIG. 1.3. Accuracy on novel duration pairs in the generalization tests that followed training. The top panel shows performance on tests that followed training with durations 0.5, 1, 2, and 4 seconds; the bottom panel shows performance following training with durations of 2, 4, 8, and 16 seconds. Data were pooled over three sessions for each pigeon. The histograms show average performance for the four birds and the lines represent standard errors of the mean. (From Fetterman & Dreyfus, 1986 by permission of Elsevier Biomedical Press B.V.)

duration reduced by some amount (a negative time-order error). This interpretation would explain the reductions in accuracy for some problems (e.g., 3 vs. 2 seconds), and the reversals (cases where accuracy was substantially below chance) in other instances (e.g., 6 vs. 4 seconds and 12 vs. 8 seconds). Second, in some instances, it appears that responding was controlled exclusively by the value of the second duration. Accuracy was consistently low on probe trials in which the longest training value was presented last, following a novel duration that was even longer (e.g., 6 vs. 4 seconds and 20 vs. 16 seconds). These results suggest that performance may have been controlled simply by the value of the second duration. During training, these values were longer than any of the other durations. Thus, the animals might have learned to respond on the basis of the single duration only. Under training conditions this strategy produced nearly perfect discrimination but led to abysmal performance on certain probe trials.

The pattern of results thus indicates that there was a temporal transposition, but that the transposition was limited by time-order errors and end effects from the longest durations.

In the final set of conditions, interstimulus intervals were inserted between the members of a duration pair. Accuracy declined as the interstimulus interval was increased from 0 to 30 seconds, but accuracy levels remained above chance even with the 30-second interstimulus interval. Accuracy measures, averaged over the four pigeons were 91, 88, 83, 72, and 72% when the interstimulus interval was 0, 2, 5, 10, and 30 seconds, respectively.

Figure 1.4 shows the effects of the interstimulus interval for the various duration pairs, and it shows accuracy at each value of the first duration (left column) or second duration (right column) collapsed across all values of the opposing pair member. For example, the filled circles on the left are accuracy scores when the 2-second first duration was paired with the 4-, 8-, or 16-second

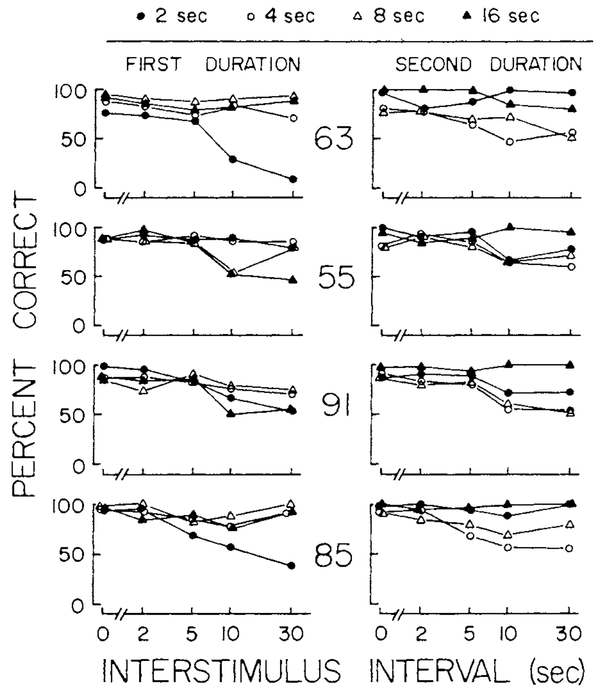


FIG. 1.4. Accuracy as a function of the interstimulus interval between the first and second durations. Accuracy is presented for each value of the first and second duration (e.g., 2 vs. 4, 8, and 16 seconds). Each point, except those for the 0-second interstimulus interval, represents performance over the final three sessions of a condition. Data for 0-second intervals are averages of two 3-day exposures. (From Fetterman & Dreyfus, 1986 by permission of Elsevier Biomedical Press B.V.)

duration. Note that the filled symbols represent extreme values (2 and 16 seconds), whereas the unfilled symbols represent intermediate values (4 and 8 seconds). The left column shows that the interstimulus interval affected performance for the various duration pairs. There was a tendency for accuracy to decline more when the first duration was extreme (i.e., 2 or 16 seconds) than when it was intermediate (i.e., 4 or 8 seconds). The more striking and informative results are found in the right column. Accuracy declined for problems in which the second member of a pair was one of the intermediate durations, approaching the chance level when the 30-second interstimulus interval was used. In contrast, the changes in accuracy were less systematic and were much less pronounced when the second duration was either 2 or 16 seconds. In fact, with a second duration of 2 or 16 seconds, there was very little change in accuracy for Pigeons 63 and 85 across the entire range of delays.

The pattern of results is not surprising but is instructive. Consider first the findings that accuracy was largely unaffected when the second duration was 2 seconds or 16 seconds. Because these were extreme values their occurrence alone would provide sufficient information for the correct response, regardless of the first duration. These values remained predictive even when the first duration was no longer remembered because of a long delay. When, however, the second duration was one of the intermediate values, either 4 or 8 seconds, the information provided by either of these two durations was not sufficient by itself; these durations were preceded sometimes by shorter and sometimes by longer durations. The findings are consistent with the transfer data in suggesting that the pigeons responded on the basis of the second duration alone whenever an extreme value was presented last. The results from the right column bear on those in the left column where accuracy tended to decrease more when the first duration was 2 or 16 seconds as opposed to 4 or 8 seconds. The extreme values of the first duration, 2 and 16 seconds, were paired more often with 4- and 8-second durations as the second-pair member. This more frequent pairing of first durations with the "confusing" second durations undoubtedly resulted in lower accuracy. Thus, changes for the first- and second-pair members appear to result from the dependencies between the values of the pair members.

Summary

Experiment 1 showed that pigeons were capable of performing a paired-comparison task involving durations. Accuracy was similar across two ranges of durations. There was some transfer to novel duration pairs on generalization tests, but transfer was limited to a certain degree by time-order errors, and by end effects that are probably inevitable when a limited number of values are used to construct stimulus pairs. Accuracy was also affected when delays were placed between the two durations, but the results showed that the pigeons could still respond appropriately even when the two durations were separated in time. The pattern of results suggests that the task was relational only insofar as the animals were forced to make relational judgments (i.e., with the intermediate durations). When, however, the animals could use the information provided by one duration only, they appeared to do so.

EXPERIMENT 2

Although the results of Experiment 1 could be interpreted as providing evidence for relational judgments in a duration comparison task, the inherent limitations that come from using few durations suggested a better procedure. Accordingly, a second experiment was designed that was identical to the first except that many duration pairs were used. For the experiment, durations were determined randomly as opposed to the fixed pairs of Experiment 1. When the first (red) duration began, a 1-second timer pulsed a probability gate set at a probability of 0.10, with the output of the probability gate ending the first duration. The same conditions held for the second (green) duration. In effect, the operation of this circuit produced durations that averaged 10 seconds, but with an actual range from 1 to 60–70 seconds. This random way of producing durations resulted in a very large number of duration pairs that, in practice, turned out to be between 600 and 700 pairs. The purpose of the experiment was to assess performance under this complex task where the opportunity to respond on the basis of a single-pair member was greatly reduced. Other than the change in the way the durations were arranged, the procedure was like that shown in Fig. 1.1 and described in the general method. The same four pigeons were the subjects.

Figure 1.5, 1.6, 1.7, and 1.8 show performance for the individual pigeons. These are matrix-type figures showing correct and incorrect responses for the different pairs of durations. Because there were so many combinations of pairs, data were pooled across four sessions for each figure. The figures only include data from durations of 20 seconds or less. Longer durations did occur, but only relatively infrequently. Because there would be few instances of these longer durations, the data were not included. Each symbol represents the outcome on an individual trial. Filled circles represent a correct response and Xs represent an incorrect response. The symbols are placed in imaginary squares that correspond to each duration pair. In Fig. 1.5, for example, there are three circles in the upper left when the first duration of 19 seconds was followed by a second duration of 1 second. The figure shows that this pair occurred three times and the animal responded correctly all three times. Similarly there was one instance of 18 seconds followed by 1 second (correct), three instances of 17 seconds followed by 1 second (all correct), and so on. The lines drawn through the matrices represent the relative differences between the durations. The 4:1 line is drawn through pairs in which the first duration was four times greater than the second

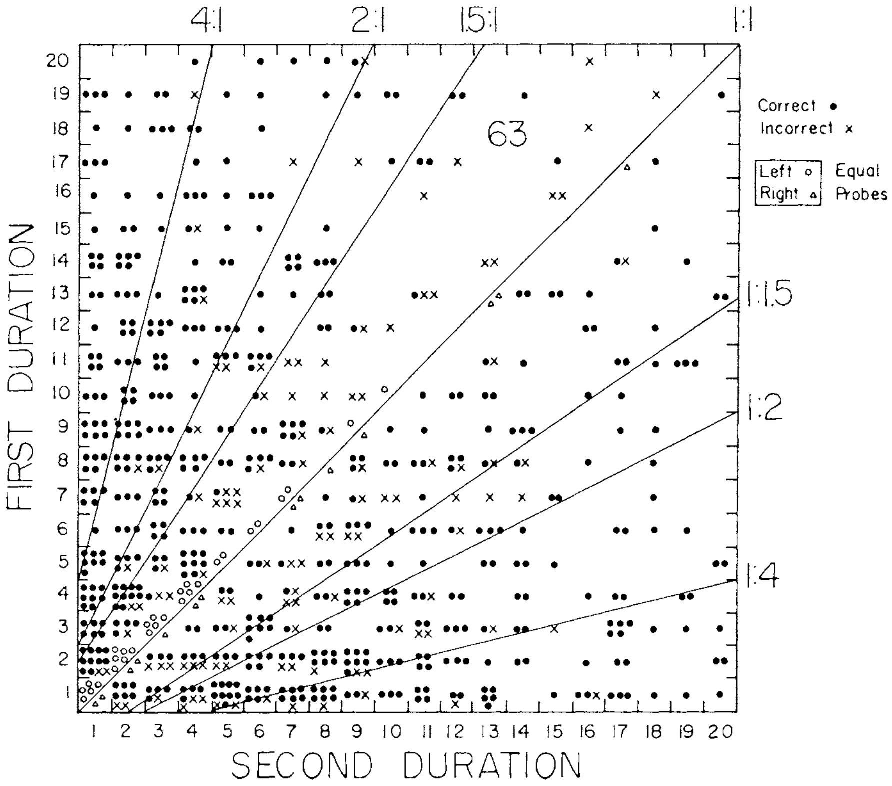


FIG. 1.5. Performance on different duration pairs for Pigeon 63. Accuracy on a particular problem is indicated by the symbols plotted at the intersection of the first and second duration. Each filled circle represents a correct choice, whereas each X represents an incorrect choice. On problems with equal first and second durations, unfilled circles and triangles represent left and right responses, respectively. Data were pooled across four sessions. Data are not included when the durations were longer than 20 seconds.

(e.g., 20 vs. 5 seconds, 16 vs. 4 seconds, 4 vs. 1 second). The 1:1 line indicates cases in which the two durations were equal. In these cases there was no correct response (and no food, only the intertrial interval). Unfilled circles indicate a left-key response, whereas unfilled triangles indicate a right-key response.

Performance was similar for the four pigeons. In all cases accuracy was high, which is indicated by the large numbers of filled circles (corrects) in each figure. Incorrect responses occurred much less often. Performance generally was similar for relatively short and relatively long durations. The major determinant of accuracy was the relative difference between the two durations. Most incorrect responses occurred when the ratios of the two durations fell between 1.5:1 and

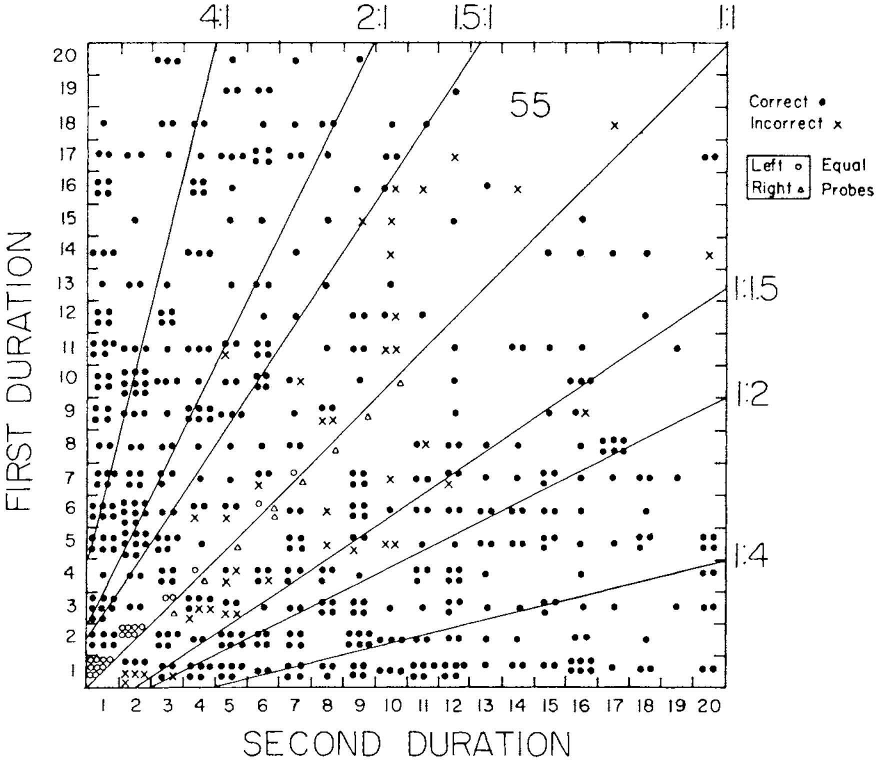


FIG. 1.6. Performance on different duration pairs for Pigeon 55. See Fig. 1.5 for description of the figure and symbols.

1:1.5. Incorrect responses occurred only rarely (and for Pigeon 55 never) when the ratios of the two durations were greater than 4:1 and 1:4. Errors became more frequent as the ratio of the two durations approached 1:1.

Figure 1.9 summarizes the individual points in Fig. 1.5–1.8. Figure 1.9 shows the probability of a right-key response (duration 1 less than duration 2) as a function of the ratio of the duration pair. The points are plotted at the midpoint of the categories of stimuli and represent the average duration pair ratio. The left column provides data for all problems. The data are characterized by ogival functions relating choice probability to duration pair ratio. The probability of a right-key response was appropriately near 0 when the first duration was relatively longer than the second (4:1 ratio). The probability increased as the durations became relatively more similar and it approached 1.0 as the second duration became progressively longer than the first (1:4 ratio).

The point of subjective equality (PSE) was calculated from the data in the left column of Fig. 1.9. The PSE represents the value of the duration-pair ratio at

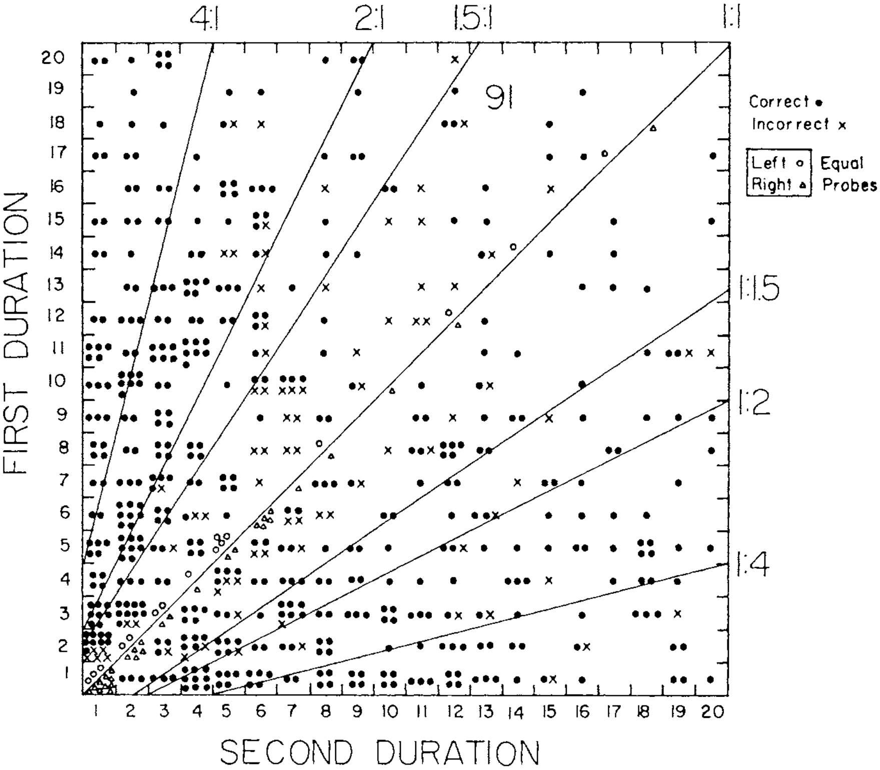


FIG. 1.7. Performance on different duration pairs for Pigeon 91. See Fig. 1.5 for description of the figure and symbols.

which the probability of a right-key response was equal to 0.50. The measure indicates whether the pigeons weighted equally the values of the first and second durations. The PSEs for the four birds averaged 1.2:1 (range 1.1:1 to 1.3:1), indicating that the two durations were judged as equal when the first duration was approximately 20% longer than the second. This finding is consistent with the negative time-order errors observed in Experiment 1.

The right column provides separate functions for cases where both durations were less than or equal to 10 seconds (filled triangles) and when one *or* both durations was longer than 10 seconds (unfilled circles). The purpose is to provide a comparison of performance when durations were relatively short and relatively long. The comparison is similar to that of Experiment 1 when two duration ranges where used. Here the wider range of durations allowed for a comparison within the same situation. The two sets of data for each bird were generally similar in appearance, with approximately equal slopes. The slope of an ogive gives an index of discrimination sensitivity. Steeper slopes reflect greater sen-

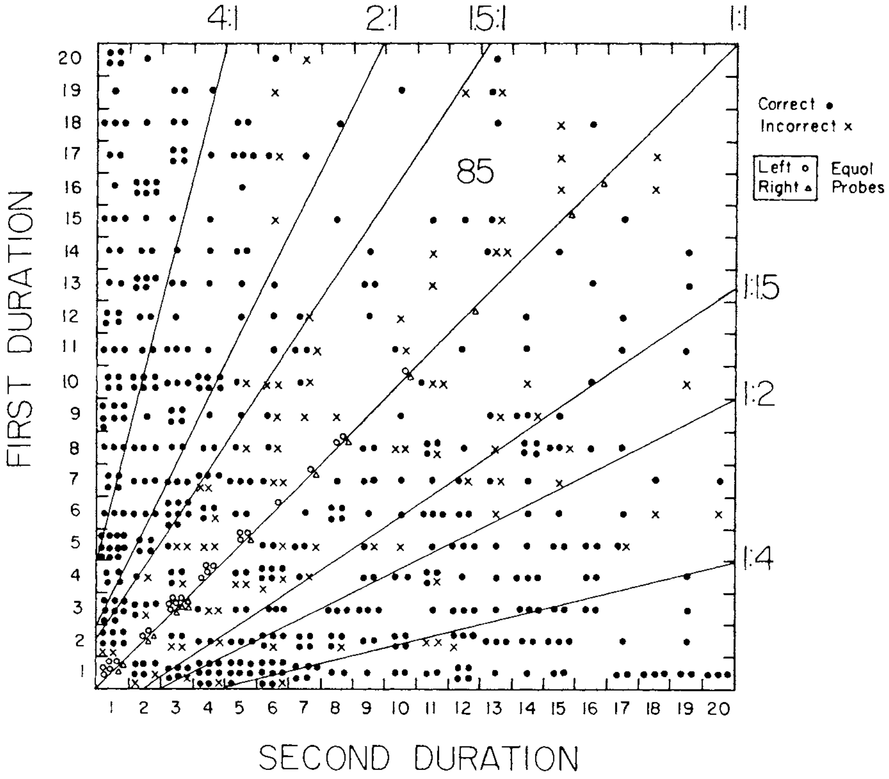


FIG. 1.8. Performance on different duration pairs for Pigeon 85. See Fig. 1.5 for description of the figure and symbols.

sitivity by showing a greater change in performance as a function of changes in the stimulus dimension. Comparable slopes indicate similar sensitivity. In the present case the finding means that discrimination performance was similar in terms of sensitivity regardless of the length of the durations.

There was one difference between the functions for three of the four pigeons. The functions for long-duration pairs were displaced to the left of the functions for short-duration pairs. Changes in the entire ogive to the left or right serve as an index of response bias (e.g., Stubbs, 1976). The bias was an increased tendency to emit a right-key response and means that there was an increased tendency to respond that the first duration was shorter than the second (or that the second was longer than the first) when one or both durations became long.

There are two possible reasons for this bias. First, when the second duration was long, the first duration was necessarily separated from a choice by a longer time than would be the case when the durations were short. This longer delay could contribute to reduced memory of the first duration. This reduced rememberance of the first duration, coupled with the long length of the second duration, might naturally produce a "biased" memory for the first duration, and thus a tendency to perceive the second duration as longer than the first. This interpretation is consistent with the occurrence of negative time-order errors described earlier. And, it is consistent with other research that has shown a bias in memory for durations when delays are imposed between the duration and choice (Spetch & Wilkie, 1983).

A second source of bias is related to the probability of occurrence of different groups of problems. Because the durations were arranged randomly, the different combinations of pairs did not occur equally, as Fig. 1.5, 1.6, 1.7, and 1.8 demonstrate. The different probabilities of occurrence of the duration pairs (e.g., long vs. short; short vs. long, etc.) could account for biased responding in a way similar to that observed for the probe trials in the first experiment. Careful examination of Fig. 1.5–1.8 reveals that when one member of a duration pair was longer than 10 seconds (all problems excluding those in the lower left quadrant) it was more probable that the other member of the pair was shorter than

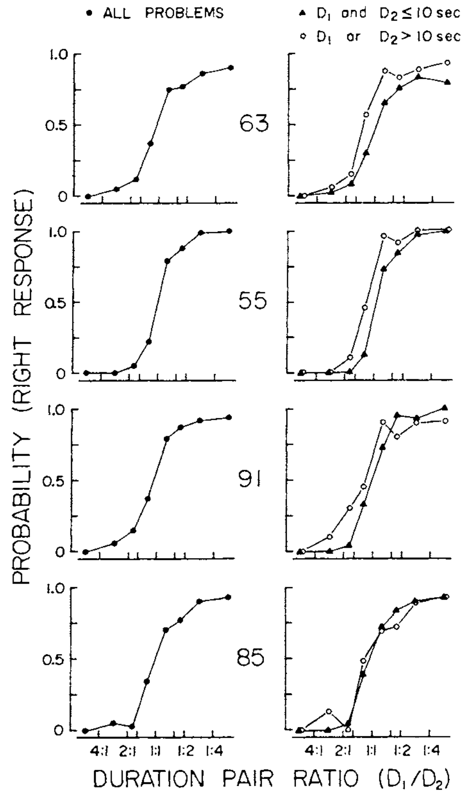


FIG. 1.9. Probability of a right-key response (reporting the second duration as longer than the first) as a function of the ratio of the duration pairs. The left column shows performance for all duration pairs. The right column shows performance when both durations were 10 seconds or less (filled triangles) and when one *or* both durations were longer than 10 seconds (unfilled circles). The points are plotted at the midpoint of the categories of stimuli and represent the average duration pair ratio for each category. The data were pooled over four sessions for each pigeon.

10 seconds (cf. the problems in the upper left and lower right quadrants with those in the upper right). Focusing upon those problems in the upper right quadrant of the figures shows that, for each animal, performance was more accurate for those problems below the 1:1 diagonal than for those above. Thus, when *both* durations were longer than 10 seconds, all birds showed a bias to respond that the second (and most recent) duration was longer. However, this same bias is not observed for problems in the lower left quadrant of Fig. 1.5–1.8, when both durations were less than 10 seconds. Perhaps bias was influenced by the probability of the different types of duration pairs. Whatever the source of bias, it should be noted that the degree of bias was not great and was not observed in all pigeons. And, in spite of a bias, sensitivity was similar over the range of durations that comprised the duration pairs.

In summary, the results of the second experiment support and extend those of the first. The pigeons were able to respond appropriately even when many different duration pairs were used. Use of many duration pairs afforded a procedural improvement that reduced end effects and forced relational judgments.

Discussion

The main findings of the experiments were the following: (a) The animals were able to discriminate the difference between two durations presented in succession; (b) discrimination accuracy depended on the relative rather than the absolute difference between the two durations; (c) performance was similar whether the durations were short or long; discrimination sensitivity was similar across the two ranges of Experiment 1 and was similar for short and long duration pairs in Experiment 2; (d) accuracy declined when a delay was interposed between the two durations, but remained relatively high until the delays approached 30 seconds; (e) negative time-order errors apparently occurred in some cases; (f) the pigeons seemed able to respond to the relational features of the task, but the occurrence of end effects suggests restrictions on discrimination based solely on stimulus relations.

The findings are consistent with previous findings on duration discrimination, and they demonstrate that the paired-comparison procedure, although different from previous procedures in a basic way, produced findings similar to other discrimination procedures that also assess time perception. There are several ways in which the different sets of data are comparable. First, the ogival functions are similar to those obtained in previous psychophysical experiments with pigeons (Stubbs, 1968) and rats (Church, 1980). Weber fractions can be computed from ogives and provide a summary measure of sensitivity (e.g., Stubbs, 1979). Weber fractions in previous psychophysical trials procedures have generally ranged around 0.20–0.30. Weber fractions for the paired-comparison task averaged .41 (range .36 to .46) for the four pigeons. The Weber fractions were obtained by dividing the difference limen (75th percentile minus the 25th percen-