

The Effective Use of Secondary Data

Russell M. Church

Brown University

In primary data analysis the individuals who collect the data also analyze it; for meta-analysis an investigator quantitatively combines the statistical results from multiple studies of a phenomenon to reach a conclusion; in secondary data analysis individuals who were not involved in the collection of the data analyze the data. Secondary data analysis may be based on the published data or it may be based on the original data. Most studies of animal cognition involve primary data analysis; it was difficult to identify any that were based on meta-analysis; secondary data analysis based on published data has been used effectively, and examples are given from the research of John Gibbon on scalar timing theory. Secondary data analysis can also be based on the original data if the original data are available in an archive. Such an archive in the field of animal cognition is feasible and desirable. © 2002 Elsevier Science (USA)

Key Words: secondary data analysis; data archives; animal cognition; primary data analysis; meta-analysis; Gibbon; scalar timing theory.

Most research in animal cognition and behavior is based upon primary data analysis in which the authors of the article collect, as well as analyze, the data. There are good reasons for this tradition: It permits the investigator to design an experiment that is most appropriate for the specific hypothesis under investigation, and it provides the investigator with direct knowledge of the conditions of the experiment and the behavior of the animals. For primary data analysis, an investigator (1) identifies a problem and an hypothesis, (2) plans an experimental design as a method to evaluate the hypothesis, (3) collects the data, (4) summarizes the data, (5) makes inferences from the

The section of this article on secondary data analysis of original data was based on a talk at the meeting of the Comparative Cognition Society, Melbourne, FL, March 17, 2000 and documents prepared for a workshop on “Data Archiving for Animal Cognition Research” sponsored by the National Institute of Mental Health and co-chaired by Russell M. Church and Howard S. Kurtzman that was held in Washington, DC, on July 19–20, 2001. The preparation of this article was supported by the National Institute of Mental Health Grant MH44234 to Brown University. Kimberly Kirkpatrick was a major contributor to the development of an archive of data of timing research: <http://www.Brown.edu/Research/Timelab>.

Address reprint requests to Russell M. Church, Department of Psychology, Box 1853, 89 Waterman St., Brown University, Providence, RI 02912. Fax: (401) 863-1300. E-mail: russell_church@brown.edu.

data, and (6) interprets the results. The integration of the experimental design and data collection stages with the data analysis and interpretation stages is the hallmark of primary data analysis.

Articles based on primary data analysis may have an important influence on further research. Any lasting impact of an article based upon primary data analysis may be estimated by citations of it in subsequent empirical articles and in reviews of the literature. The earlier article may be cited for its statement of the problem, its methods, its published results, or its conclusions. The subsequent articles rarely involve any further analyses of the original data used for the published results.

In “secondary data analysis,” the individual or group that analyzes the data is not involved in the planning of the experiment or the collection of the data. Such analysis can be done based upon information that is available in the statistical information in the published articles, the data available in the text, tables, graphs, and appendices of the published articles, or upon the original data.

META-ANALYSIS

Meta-analysis refers to a quantitative combination of the statistical information from multiple studies of a given phenomenon. It provides a rigorous way to summarize the results of these studies. An excellent description of the approach is provided by Mullen (1989). The bases for meta-analysis were developed by R. A. Fisher and others in the 1920s and 1930s (Fisher, 1938); even the way to combine probabilities from independent studies was well known to researchers (Mosteller & Bush, 1954) long before the term “meta-analysis” was coined by Gene Glass (1976). The method was used to combine the results of 375 controlled evaluations of psychotherapy and counseling to reach the conclusions that therapy worked and that there were few important differences in the effectiveness of different types of therapy (Smith & Glass, 1977). A search of PsychInfo from January 1887 through November 2000 identified 3457 meta-analysis studies—all since 1977. Although this procedure has been used extensively in other areas of psychology, a search of PsychInfo did not identify any meta-analysis studies in *Animal Learning & Behavior*, *Behavioural Processes*, *Journal of Comparative Psychology*, *Journal of Experimental Psychology: Animal Behavior Processes*, *Learning and Motivation*, or the *Quarterly Journal of Experimental Psychology (B)*. Three meta-analysis studies were published in *Animal Behaviour* and one in *Behaviour*, but the only one related to animal cognition and behavior was about the role of magnetoreception in human navigation (Baker, 1987). There has also been a meta-analysis of the difference between prospective and retrospective time estimations of human participants (Block & Zakay, 1997). One meta-analysis article was published in *Journal of the Experimental Analysis of Behavior*, according to the indexing software (McSweeney, Farmer, Dougan, & Whipple, 1986). This was an extensive

review of the quantitative results related to the applicability of the generalized matching law to the results of experiments on multiple schedules of reinforcement. The authors did not refer to this as a meta-analysis study in their title, abstract, or key words, and, with the definitions used in this article, it should be classified as an excellent example of secondary data analysis of published data. A meta-analysis has been done of field studies relating depletion of resource patches to initial resource density (Dolman & Sutherland, 1997). It is not clear why meta-analysis has been rarely, if ever, used in research on animal cognition and behavior. Although there are problems with obtaining a random sample of studies, establishing independence of observations, and determining the comparability of the conditions of the studies, the problems are not unique to this field. Probably the lack of use of meta-analysis of studies of animal cognition and behavior is due to research conventions in the field. Whether or not meta-analysis would provide useful quantitative measures to supplement narrative reviews has not been evaluated. An essential weakness of meta-analysis is that it relies upon the statistical analysis of published data and thus lacks the versatility that is possible in an examination of the raw data.

SECONDARY ANALYSIS OF PUBLISHED RESULTS

In both his research and theoretical articles, John Gibbon made effective use of secondary analysis of published results. He first described "scalar timing" in an article in the *Journal of Mathematical Psychology* (Gibbon, 1971). In one figure he demonstrated that the latency of avoidance responding is a linear function of the warning signal duration (see Fig. 1). He used the reported results of four experiments (Anderson, 1969; Hyman, 1969; Kamin, 1954; Low & Low, 1962), redrew the functions on a common scale, and edited the data in one case by eliminating the data from one animal that showed strong order effects. He interpreted the linearity of the function in terms of scalar timing, and the positive intercept in terms of "motor time in executing the response." However, he noted that the form of the interresponse time distributions provided a much more stringent test. This is presumably because there are many more ways to produce a linear relationship between mean latency and stimulus duration than there are to produce an identity of multiple functions from results obtained under different conditions.

In the analysis of data from Sidman avoidance procedures in his laboratory and in two other laboratories (Anger, 1963; Verhave, 1959), Gibbon (1971) showed that interresponse time distributions for single rats at different response-shock intervals were essentially of the same type when time was scaled in relative units. This is an example of what was later called superposition and time scale invariance. Because the published data from the various experiments reported different dependent variables, Gibbon found it necessary to show the linearity of the latency with stimulus duration on the

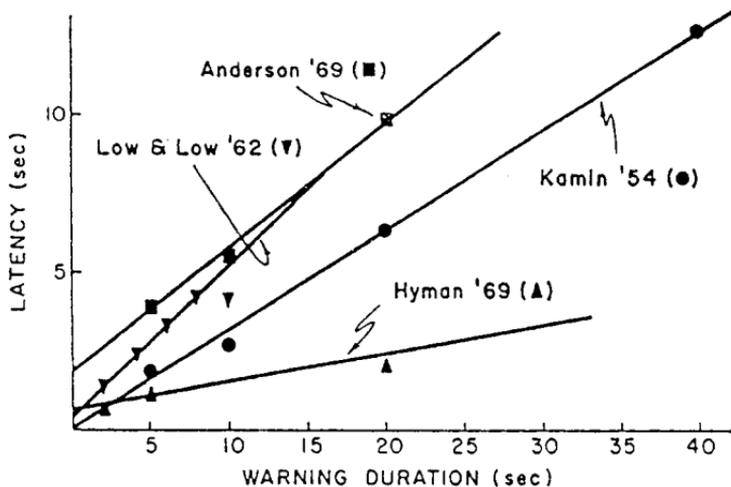


FIG. 1. Latency of avoidance responding as a function of warning signal duration. t is time since the interval began; T is the mixed interval value; and M is a fixed latency to begin timing. The figure is reprinted from Gibbon (1971). It was based on Gibbon's secondary analysis of data from Anderson (1969), Low and Low (1962), Kamin (1954), and Hyman (1969).

basis of one set of experiments and the superposition effect on another set of experiments.

Gibbon (1971) derived explicit solutions for a model of the mean inter-response and intershock time functions for several free-operant avoidance schedules. In one figure he showed that the model provided excellent fits to the data from five individual animals from four different experiments (Clark & Hull, 1966; Hake, 1968; Sidman, 1953; Verhave, 1959). Although these experiments involved different species (rats, dogs, and monkeys) and different procedures, Gibbon demonstrated that simple quantitative functions based on scalar timing applied to results from all of them. This analysis was extended to the effect of amount of reduction in shock density in a theoretical article that included data from published experiments of others (Gibbon, 1972). An unwritten message from this article is that it is not necessary to design a specific experiment to determine whether or not a quantitative principle is applicable. Confidence in the generality of a principle may be increased by the range of studies to which it applies and the fact that the author did not design the experiment for the purpose of illustrating the principle.

According to the Science Citation Index, this important article (Gibbon, 1971) in the *Journal of Mathematical Psychology* has been cited only 21 times, and only four times in the past decade. In an influential article in the *Psychological Review* entitled "Scalar expectancy theory and Weber's law in animal timing" (Gibbon, 1977), scalar timing was applied to a wider range of procedures. This article has been cited 358 times, and the rate of citations

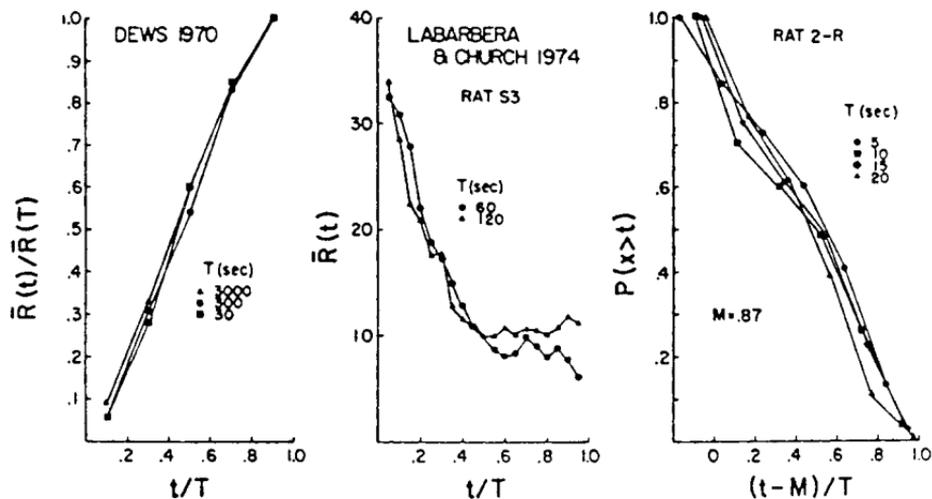


FIG. 2. Three examples of relativistic timing. This figure is reprinted from Gibbon (1977). It was based on secondary analysis of published data from Dews (1970), LaBarbera and Church (1974), and unpublished data from Gibbon's laboratory.

does not appear to be decreasing: Over a third of the citations (126) were in the past 5 years (1996 through 2000).

The first figure in this theoretical article (Gibbon, 1977) demonstrated the relativistic nature of timing behavior (Fig. 2). The three figures came from unpublished research from Gibbon's laboratory and two other experiments that used quite different methods: pigeons or rats, fixed or variable interval schedules of reinforcement, various ranges of intervals, and food or shock reinforcers (Dews, 1970; LaBarbera & Church, 1974). The function forms obtained in the three experiments are different (increasing or decreasing, linear or nonlinear), and different independent and dependent variables were used. The one regularity was that all the functions within each of the experiments were essentially the same. This required that the independent variable represent time in relative rather than in absolute units.

In other figures in this article, Gibbon (1977) showed that the mean and standard deviation increased linearly with the fixed interval, that the coefficient of variation (the standard deviation divided by the mean) was approximately constant (Schneider, 1969; Schneider & Neuringer, 1972) in a fixed time schedule (Killeen, 1975), that the mean time to the second response in a progressive interval schedule was linearly related to the interval schedule on a log-log scale (Harzem, 1969), that the function relating normalized activity to relative time was the same at three different intervals between food presentations (Killeen, 1975), and that the function relating probability of a "short" response to relative time was the same for various ranges of time intervals (Stubbs, 1968).

In the application of scalar expectancy theory to choice, Gibbon (1977)

used data from Chung (1965), Chung and Herrnstein (1967), Duncan and Fantino (1970), Killeen (1970), Moffitt and Shimp (1971), Rachlin and Green (1972), Shimp (1968, 1969), and Staddon (1968) to compare the predictions of scalar expectancy theory with matching. The quantitative predictions of scalar expectancy theory were excellent in all five of the figures; the matching predictions were not. Gibbon's 1997 article provided broad support for the general principles of scalar expectancy theory and for the application of Weber's law in animal timing.

The location of the data points in the published articles were carefully measured by enlarging the figures photographically to 8×10 inches, and using a Gerber scientific GraphAnalogue ruler (Model GA-103) to measure the location within 1% of the data values (Gibbon, 1977, footnote 2).¹ Other approaches are to make an overhead transparency of a graph and project it on large graph paper on a wall, or to use a copy machine to enlarge a graph and enter the observed values in a spreadsheet (in millimeters), calculate the slope and intercept of the scale transforms, and transform the units from lengths to units of the dependent variable. A modern approach is to use a scanner to digitize a graph that can be saved in an appropriate format and then use software (such as DataThief, from www.shareware.com) to determine the location of the points. These procedures provide ways to obtain reasonably accurate estimates of the values shown in graphs. The main limitation is that the dependent variables displayed in most graphs represent highly summarized information and not the original data. In some cases, such as Killeen (1975), John Gibbon requested and received copies of original data sheets and documentation.

In many of his studies of avoidance of shock, autoshaping, and interval schedules of positive reinforcement, and choice, Gibbon supported the generality of conclusions by using published quantitative data from other experiments. The figure that included data from the largest number of other experiments showed that the number of reinforcers to acquisition of an autoshaped response by pigeons was linearly related to the ratio of the intertrial duration

¹ Information about the Gerber GraphAnalogue is available at <http://www.gerberscientific.com>. It includes the following story from 1945:

While studying aeronautical engineering at Rensselaer Polytechnic Institute, H. Joseph Gerber was frequently required to solve intricate and time-consuming mathematical problems, many involving the plotting of points. One night, in an effort to reduce the tedious and repetitive nature of his homework assignments, he created an "expandable ruler" from the elastic waistband of his pajama bottoms—thereby inventing a new method of scaling distances between points. An instant sensation on campus, his invention was eventually modified to incorporate a unique triangular spring . . . giving birth to the Gerber Variable Scale®. This handheld device, hailed as the greatest invention since the slide rule, is still in production today—more than 50 years after the fact—with thousands in use around the world. And, it is on permanent display at the Smithsonian Institution's Museum of American History.

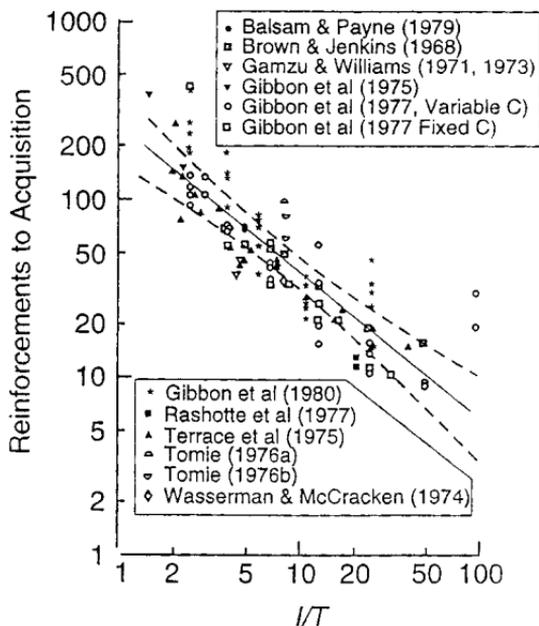


FIG. 3. Number of reinforcers to an acquisition criterion as a function of the ratio of intertrial duration (I) to stimulus duration (T). The figure is reprinted from Gallistel and Gibbon (2000), which was redrawn from Gibbon and Balsam (1981). It was based on secondary analysis from 12 published experiments listed in the legend.

to trial duration on a log-log scale (Gibbon & Balsam, 1981) (see Fig. 3). In autoshaping, food may be presented at some interval (I) after the previous food, and the stimulus may begin at some interval (T) prior to the food. For example, I may be 100 s, and T may be 10 s. This makes the I/T ratio 10, a ratio that produces a moderate acquisition score.

In his description of the historical and causal origins of scalar timing theory, Gibbon (1991) replotted data from a well-known study of fixed interval responding that demonstrated that the response rate on individual intervals abruptly changed from a low rate to a high rate and that this break point varied from one interval to another (Schneider, 1969). The top panel of Figure 4 shows the conventional representation of mean response rate as a function of time; the bottom panel shows that both the mean and the standard deviation of the breakpoint increased linearly with the interval duration, although at different slopes. The linear functions accounted for about 99% of the variance. This is one of many cases in which Gibbon's analysis went beyond the original published report. Schneider had demonstrated the linearity of the mean of the breakpoint as a function of the fixed interval, and he provided enough data for Gibbon to extend this to demonstrate the linearity of the standard deviation.

John Gibbon made extensive use of secondary data analysis in the analysis

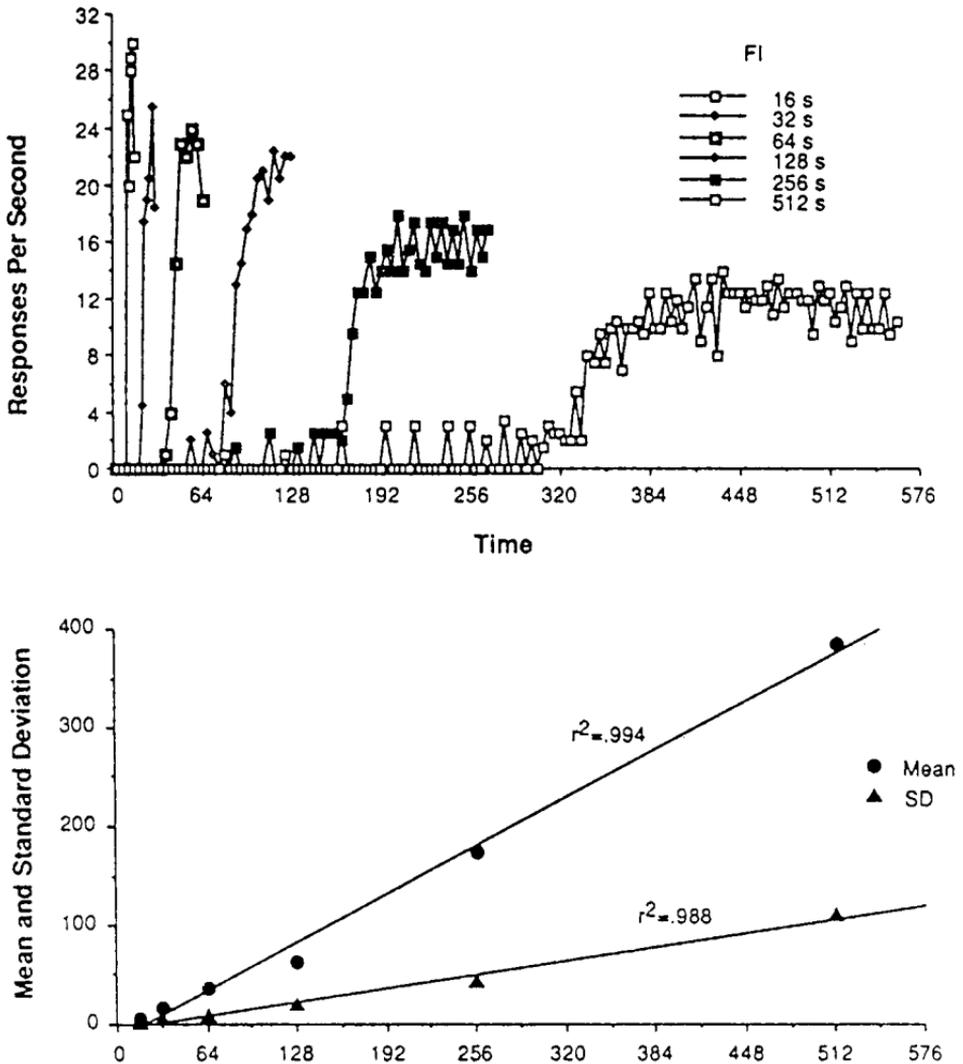


FIG. 4. Response rate as a function of time since food delivery on a fixed interval schedule of reinforcement (top); mean and standard deviation of the break point as a function of the fixed interval (bottom). The figure is reprinted from Gibbon (1991). It was based on secondary analysis from published data from Schneider (1969).

of human as well as animal timing. In one figure of a review article on the neurobiology of temporal cognition (Gibbon, Malapani, Dale, & Gallistel, 1997), the coefficient of variation was shown as a function of the duration of the interval being timed, from under 100 ms to over 2 h, based on 28 studies with humans and 15 studies with animals.

In their major statement regarding the relationship of timing and conditioning, Gallistel and Gibbon (2000) included 30 figures: over half of them in-

volved secondary data analysis of published data; the others involved primary data analysis, or a specification of model or procedure. The most important feature of secondary data analysis is that the actual quantitative results of published research is taken seriously. This can be used to establish the generality of a quantitative function, and it can also be used to identify problems of theoretical interest. The close examination of the quantitative results of published studies has often facilitated the development and testing of new quantitative theory: An excellent example of this is Killeen's article in this issue.

SECONDARY ANALYSIS OF ORIGINAL DATA

For a primary analysis, the investigator must select particular summary measures to report in the text, figures, tables, and short appendices. The decision about how to summarize the data is an important one, because it is irreversible. It is seldom possible to regenerate the original data from the summaries and, therefore, the published summary measures of performance can only rarely be used to examine alternative measures of performance. For example, a study of classical conditioning that reports absolute or relative rates (or probabilities) of responding in the presence and absence of a stimulus cannot be used to evaluate a real-time theory of conditioning because the time of responding has been eliminated from the record.

The major limitation of secondary data analysis of the published data is that, due to the constraints of space in paper journals, the publication includes only a summarized version of the original data. There is considerable information in the original data that cannot be recovered from the summary measures reported in the published article. If the original data were available, this limitation of secondary data analysis would disappear. Some journals, such as the *Journal of the Experimental Analysis of Behavior*, have a long record of publishing large tables of data, but most rarely have space for more than a few highly summarized tables.

Figures and highly summarized tables seldom provide the necessary information for the analysis of alternative dependent variables or the examination of different problems. For example, Crystal, Church, and Broadbent (1997) reported results of several experiments to examine systematic nonlinearities in rats' memory for time. The published analyses based upon a particular dependent variable, the "start time," could not be used by an analyst who wanted to examine a different dependent variable, such as "time of median response." From the original data, however, it is possible to examine this dependent variable and a different problem, such as the scalar nature of the behavior. Figure 5 shows that both the mean and the standard deviation of the time of the median response of rats in 66 different interval durations between 10 and 140 s is approximately linear; therefore, the ratio of the standard deviation to the mean (the coefficient of variation) would also be approximately linear. This is an illustration of the use of data from an experi-

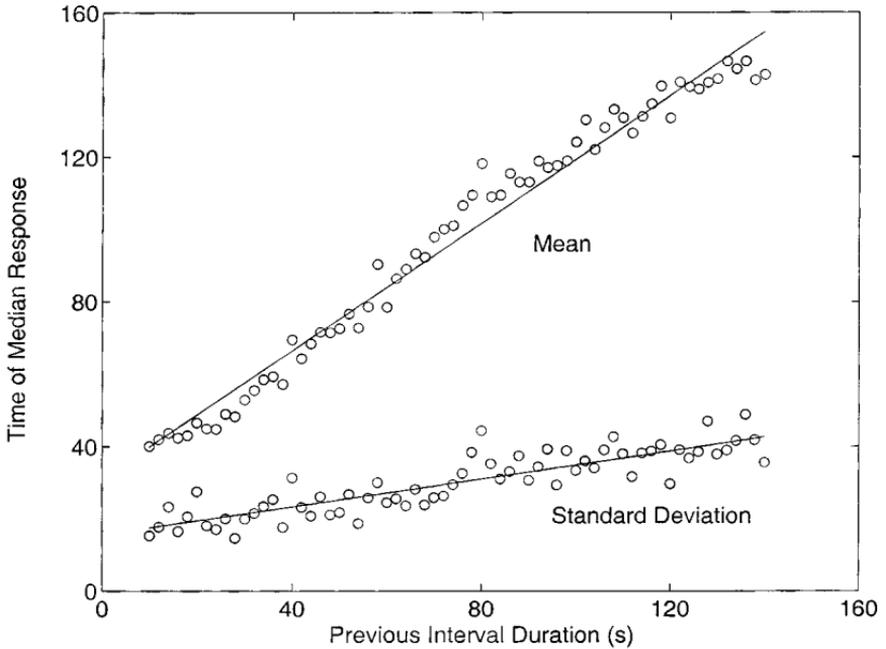


FIG. 5. Mean and standard deviation of the time of the median response as a function of previous interval duration. This figure is based on secondary data analysis of the original data from Group 2 of Experiment 2 of Crystal, Church, and Broadbent (1997) on the web site <http://www.brown.edu/Research/Timelab>.

ment designed for the study of one topic (systematic nonlinearities in timing) for obtaining evidence on a different topic (Weber's law for timing).

Every year research psychologists collect an enormous amount of data. They summarize it, they attempt to explain it, and they publish articles based on the research. The original data may be in notebooks, in stacks of data forms and, increasingly, in computer files. Various guidelines for data retention and data sharing have been made by journal editors, by the ethical guidelines for members of the American Psychological Association, and others. Many psychologists maintain these records for five years or more, but it currently is difficult to analyze them further: The details of the data formats may be difficult to reconstruct, and they may be stored on media that are no longer standard. Some investigators are willing and able to make their original data available upon request to others, but, for all practical purposes, after publication most original data are inaccessible. At present, any influence of a research project must be due to the features that can be published, such as the procedure, the summarized data, or the conclusions; it cannot be based on the original data.

Many experimenters currently record extensive data from experiments, and these records may be saved on removable disks or other storage media.

For example, the experimenter may automatically record the times of occurrence of several stimuli and responses during many sessions. These are more complete than more summarized measures which do not record the specific times of occurrence of each measured event, and these data provide more information than is usually available in the published articles. For example, they would provide information about each animal on each session. Secondary data analysis could be done on the original data, if the original data were available.

It is now feasible to make original data available on the internet. At present some investigators have made data available on personal websites, and this is reasonably easy to do. The original data from many experiments can be represented as the times of occurrence of different events. These can be represented in text format in a table with many rows and two columns. The original data from other experiments can be represented as numbers or words in text format in tables with many rows and more than two columns. For some purposes, such as archiving of stimuli, other formats are required. Scientific societies or other organizations could provide an important service by facilitating the development of standard format conventions and by providing storage space for data that could be mirrored at other sites.

The archive should be easy to use: It should be well organized and compact; it should use standard formats; and the software should be written in a way to guarantee fast, reliable transmission. Most critically, each data set in the archive should be accompanied by good documentation that specifies completely and in a standardized form the experimental procedures and the codes used. Sole reliance on personal archives is not a sufficient long-term solution to providing permanent and wide access to original data. A useful archive of data should contain data from research on a topic that is important and focused (e.g., animal cognition), but that is broader than the research of a single laboratory.

A data archive should contain a large quantity of high-quality data with both hierarchical organization of subtopics and features to facilitate searches. An archive of data would be particularly useful if it contained various forms of assistance, such as references to the published literature, links to other relevant sites, data analysis tools, theory tools, and technical support. Data sets linked to full-text versions of published articles that have gone through the peer review process are particularly valuable, especially if the full-text versions are widely available.

A major concern is that investigators will not put their data in archives. They may consider the effort of putting the data into a standard format to outweigh the value of the general accessibility of the data. Many articles are rarely cited, and may be rarely read; many data sets may never be used for secondary data analysis. Some investigators may also be concerned that errors will be exposed or that others will publish articles based upon their data or will propose alternative explanations of their data. Examples in which

further analysis of data led to the identification of errors include Church, Crystal, and Collyer (1996), Gaffan and Gaffan (1992), and LoLordo and Ross (1991).

There are many reasons investigators might want to put data in an archive. The scholarly values include the advancement of knowledge, the identification of errors, the announcement of negative results, and the support and amplification of a published article. The societal values include the maximization of information available to everyone and the minimization of the number of animals required. There are even reasons of self-interest to investigators to put data in an archive: It could increase the number of citations and increase the credibility of the data.

Archives of original data might be used by empiricists, theorists, and scholars. Empiricists might use the archives for further analyses, improvement of experimental designs, and detection of errors in previous research. Theorists might use them for development of quantitative models and for tests of current quantitative models. Many different types of scholars, who are now excluded from participation in research on animal cognition, might use them. These include faculty without laboratory resources, teachers, students, writers of in-depth review articles, data analysts in other fields, and gifted amateurs.

Secondary analysis of the original data is not currently being used in the field of animal cognition, although it is being used extensively in other fields, such as astronomy, high-energy physics, the genome project, and psychological health surveys. It has been used primarily in fields in which there is a large amount of data and the cost of the data collection is particularly high. Animal cognition experiments are often expensive because they may require specialized laboratory facilities for the maintenance of animals and for the experimental control and recording of behavior, and many of the experiments require lengthy training protocols. Thus, the data in animal cognition may be an excellent candidate for secondary data analysis. Although this may require some adjustment of editorial policy and attitudes of investigators and peers toward the analysis of data collected by others, such an adjustment could facilitate progress in the field. Data archives may become a resource as useful as a library, and they are feasible to develop.

REFERENCES

- Anderson, N. H. (1969). Variation of CS-US interval in long-term avoidance conditioning in the rat with wheel turn and with shuttle tasks. *Journal of Comparative and Physiological Psychology*, **68**, 100-106.
- Anger, D. (1963). The role of temporal discriminations in the reinforcement of Sidman avoidance behavior. *Journal of the Experimental Analysis of Behavior* (Suppl.), 477-506.
- Baker, R. R. (1988). Human navigation and magnetoreception: The Manchester experiments do replicate. *Animal Behaviour*, **35**, 691-704.
- Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytical review. *Psychonomic Bulletin & Review*, **4**, 184-197.

- Chung, S. (1965). Effects of delayed reinforcement in a concurrent situation. *Journal of the Experimental Analysis of Behavior*, **8**, 439–444.
- Chung, S., & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, **10**, 67–74.
- Church, R. M., Crystal, J. D., & Collyer, C. E. (1996). Correction of errors in scientific research. *Behavior Research Methods, Instruments, & Computers*, **28**, 305–310.
- Clark, F. C., & Hull, L. D. (1966). Free-operant avoidance as a function of the response-shock = shock-shock interval. *Journal of the Experimental Analysis of Behavior*, **9**, 641–647.
- Crystal, J. D., Church, R. M., & Broadbent, H. A. (1997). Systematic nonlinearities in the memory representation of time. *Journal of Experimental Psychology: Animal Behavior Processes*, **23**, 267–282.
- Dews, P. B. (1970). The theory of fixed-interval responding. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 43–61). New York: Appleton-Century-Crofts.
- Dolman, P. M., & Sutherland, W. J. (1997). Spatial patterns of depletion imposed by foraging vertebrates: Theory, review and meta-analysis. *Journal of Animal Ecology*, **66**, 481–494.
- Duncan, B., & Fantino, E. (1970). Choice for periodic schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, **14**, 73–86.
- Fisher, R. A. (1938). *Statistical methods for research workers* (7th ed.). London: Oliver and Boyd.
- Gaffan, E. A., & Gaffan, D. (1992). Less-than-expected variability in evidence for primacy and von Restorff effects in rats' nonspatial memory. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 298–301.
- Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, **107**, 289–344.
- Gibbon, J. (1971). Scalar timing and semi-Markov chains in free-operant avoidance. *Journal of Mathematical Psychology*, **8**, 109–138.
- Gibbon, J. (1972). Timing and discrimination of shock density in avoidance. *Psychological Review*, **79**, 68–92.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, **84**, 279–325.
- Gibbon, J. (1991). Origins of scalar timing. *Learning and Motivation*, **22**, 3–38.
- Gibbon, J., & Balsam, P. (1981). Spreading association in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219–253). New York: Academic Press.
- Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition. *Advances and challenges. Current Opinion in Neurobiology*, **7**, 170–184.
- Glass, G. (1976). Primary, secondary and meta-analysis of research. *Educational Research*, **5**, 3–8.
- Hake, D. F. (1968). Actual versus potential shock in making shock situations function as negative reinforcers. *Journal of the Experimental Analysis of Behavior*, **1**, 385–403.
- Harzem, P. (1969). Temporal discrimination. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning*. New York: Academic Press.
- Hyman, A. (1969). Two temporal parameters of free-operant discriminated avoidance in the rhesus monkey. *Journal of the Experimental Analysis of Behavior*, **12**, 641–648.
- Kamin, L. J. (1954). Traumatic avoidance learning: the effects of CS–US interval with a trace-

- conditioning procedure. *Journal of Comparative and Physiological Psychology*, **47**, 65–72.
- Killeen, P. (1970). Preference for fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, **14**, 117–124.
- Killeen, P. (1975). On the temporal control of behavior. *Psychological Review*, **82**, 89–115.
- LaBarbera, J. D., & Church, R. M. (1974). Magnitude of fear as a function of the expected time to an aversive event. *Animal Learning and Behavior*, **2**, 199–202.
- LoLordo, V. M., & Ross, R. T. (1991). Retraction of Ross and LoLordo findings concerning blocking in serial feature-positive discriminations. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 402–406.
- Low, L. A., & Low, H. I. (1962). Effects of CS–US interval length upon avoidance responding. *Journal of Comparative and Physiological Psychology*, **55**, 1059–1061.
- McSweeney, F. K., Farmer, V. A., Dougan, J. D., & Whipple, J. E. (1986). The generalized matching law as a description of multiple-schedule responding. *Journal of the Experimental Analysis of Behavior*, **45**, 83–101.
- Moffitt, M., & Shimp, C. P. (1971). Two-key concurrent paced variable-interval paced variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, **16**, 39–49.
- Mosteller, F., and Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology* (pp. 289–334). Cambridge: Addison-Wesley.
- Mullen, B. (1989). *Advanced BASIC meta-analysis*. Hillsdale, NJ: Erlbaum.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, **17**, 15–22.
- Schneider, B. A. (1969). A two-state analysis of fixed-interval responding in the pigeon. *Journal of the Experimental Analysis of Behavior*, **12**, 677–688.
- Schneider, B. A., & Neuringer, A. J. (1972). Responding under discrete-trial fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, **18**, 187–199.
- Shimp, C. P. (1968). Magnitude of frequency of reinforcement and frequencies of interresponse times. *Journal of the Experimental Analysis of Behavior*, **11**, 525–535.
- Shimp, C. P. (1969). The concurrent reinforcement of two interresponse times: The relative frequency of an interresponse time equals its relative harmonic length. *Journal of the Experimental Analysis of Behavior*, **12**, 403–411.
- Sidman, M. (1953). Two temporal parameters of the maintenance of avoidance behavior by the white rat. *Journal of Comparative and Physiological Psychology*, **46**, 253–261.
- Smith, M. L., & Glass, G. (1977). Meta-analysis of psychotherapy outcomes. *American Psychologist*, **32**, 752–760.
- Staddon, J. E. R. (1968). Spaced responding and choice: A preliminary analysis. *Journal of the Experimental Analysis of Behavior*, **11**, 669–682.
- Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, **11**, 223–238.
- Verhave, T. (1959). Avoidance responding as a function of simultaneous and equal changes in two temporal parameters. *Journal of the Experimental Analysis of Behavior*, **2**, 185–190.