

SERENDIPITY IN ANIMAL EXPERIMENTATION: EXAMPLES FROM DURATION SCALING IN RATS

Hannes Eisler

University of Stockholm, Stockholm, Sweden

ABSTRACT: In the scaling experiment proper, 8 rats had to reproduce 10 randomly presented time intervals ranging from 1.3 to 20 s. The beginning of the reproduction was separated from the end of the standard by a 300 ms interruption of the sound indicating the durations. The rat determined the length of the reproduction by pressing a lever, thereby terminating the sound. The scaling model and the final result of the experiment are briefly described. However, before this final phase of the experiment, the rats had to learn to attend to the interruption. This was achieved in Phase 2 of the experiment by defining lever presses during presentation of the standard as incorrect behavior, which was penalized by withholding the reward and lengthening the actual standard duration by 4 s. Scrutiny of the Phase 2 data revealed two unexpected—serendipitous—findings. 1. The rats learned that they had committed an error (a faulty lever press during the standard) before they learned to suspend their lever presses; shorter response latencies for to-be-rewarded than for not-to-be-rewarded trials clearly indicated: “Knowing before doing.” 2. A study of the distribution of first lever presses during the standard showed (a) that these lever presses were not evenly distributed (the hypothesis of a negative exponential distribution was rejected), and, more interestingly, (b) that 7 of the 8 rats hardly ever pressed the lever during a certain interval (for 5 of the rats the interval 3.3-4.5 s): “Temporal holes in the latency distributions.” It pays to look not only at data when learning has been accomplished, but also during acquisition!

In a course I teach on “How to write papers for psychological journals,” one of the rules presented in slogan form is “Don’t write an autobiography.” However, in order to describe serendipitous findings, as well as fortuitous events influencing the particular research I shall relate, namely duration scaling in rats, I have to break this rule.

The soil in which the two roots of the study grew was the interest of this Department of Psychology, under the direction of Gösta Ekman, in scaling in general, and in studies of time perception in particular. One of the still prevalent issues in direct scaling methods like magnitude estimation or ratio setting is to what extent the scales obtained mirror observers’ number behavior rather than their sensations. I thought that the role played by number behavior in rats

Address correspondence to Dr. Hannes Eisler, Department of Psychology, University of Stockholm, S-106 91 Stockholm, Sweden.

might be markedly smaller than in human observers, so that scaling in rats, if yielding results similar to those from psychology students, would support the validity of subjective scales obtained by direct methods. This idea was one of the two roots. The other was more concrete; at a talk at a congress in New York City, timing experiments with rats were described in which drug effects were studied. The outcome was interpreted as changes in activity levels and I thought that time or duration scales could give a more precise explanation. At that time I was working at Harvard University (with S. S. Stevens), where I had access to rats and instrumental conditioning equipment—a lucky coincidence.

I designed the experiment—I shall mention the details later—and got some help with the wiring of the rack. (This was before computers entered laboratories.) I knew very little about relays and the other devices used, and practically nothing about rats, except that they should be kept at 80% of their free feeding body weight when used in experiments. I got four rats and had learned the following when I left Harvard, long before the experiment was completed: 1. Female rats are not very good experimental animals (all four were females); 2. There are large individual differences; 3. Rats grow throughout their life. I had kept the poor animals at 80% of their weight when I took over and when I left Harvard they were emaciated; my departure probably saved them from death by starvation.

Back in Stockholm I made only two more mistakes: I underestimated the complexity of the apparatus required for my experiment and I tried to use mice (which were very cheap) as experimental animals. I had two apparatuses built that did not work reliably and the mice never seemed to learn, unlike the Harvard rats—a typical example of confounded variables. Were the mice or the apparatus designers to blame? Eventually the department hired a new and very capable engineer and I purchased male blackhooded rats from Denmark.

The scaling was to build on duration reproduction data. Ten different standard durations, ranging from 1.3 to 20 s in equal log steps, were presented at random as a distorted tone. After an interruption of 300 ms the sound resumed and was shut off by a lever press. The rats' (final) task was to press the lever when they experienced the duration of the sound after the interruption as equal to the standard duration presented in that trial (see below). Eight rats were included in the experiment.

THE PARALLEL-CLOCK MODEL

I had certain—not too explicit—ideas about how duration reproduction data could answer the questions posed: that a duration scale

of the level obtainable would give a better description of drug effects than the level of activity and that psychophysical scales are not an artefact of number behavior. Anyhow, in order to get an impression of the working of the experiment, to check on the suitability of the chosen stimuli as well as of the length of the interruption, I carried out a pilot study with human observers (i.e., students of psychology) as models for the rats. The outcome was unexpected. The data indicated that the observers did not keep the standard durations in memory. Instead they used two "sensory registers," one accumulating the total subjective duration (the subjective duration corresponding to sum of the standard and the reproduced durations) and the other the subjective duration corresponding to the reproduced duration. The two durations, standard and reproduced, are experienced as equally long when the difference between the contents of the two registers equals the contents of the second register (Eisler, 1975). Figure 1 should make this idea clear. This model of duration reproduction behavior (the "parallel-clock model") can be said, from the researcher's point of view, to be tantamount to halving the total duration. Regarding the reproduced duration as subjectively half the total duration makes it possible to construct a subjective ratio scale for duration.

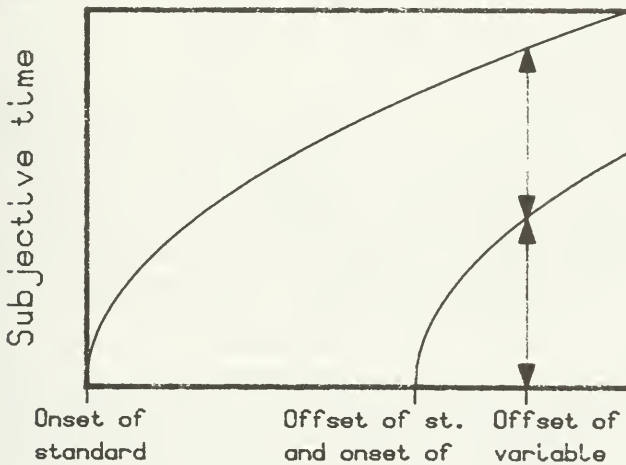


FIGURE 1. Duration reproduction according to the parallel-clock model. Subjective versus total physical duration (*left curve*) and versus reproduced duration (*right curve*). When the difference between these two subjective durations (*upper arrow*) equals the subjective reproduced duration (*lower arrow*), the observer reports equality between standard and reproduction by shutting off the sound.

Denoting subjective reproduced duration Ψ_r and subjective total duration Ψ_t , we obtain from the parallel-clock model

$$\Psi_r = (1/2) \Psi_t \quad (1)$$

Inserting Stevens' psychophysical power law

$$\Psi = (\Phi - \Phi_0)^\beta \quad (2)$$

where Ψ denotes subjective duration, Φ clock duration, and β and Φ_0 are constants, into Equation 1, we obtain

$$(\Phi_r - \Phi_0)^\beta = (1/2)(\Phi_s + \Phi_r - \Phi_0)^\beta \quad (3)$$

The subscript *r* refers to reproduction and *s* to standard; the total clock time, to which the total subjective duration Ψ_t corresponds, is $\Phi_s + \Phi_r$. By rearranging Equation 3 it can be seen that the reproduced duration is a linear function of the total duration. The parameters of the psychophysical function can be computed from the slope and intercept of a plot of reproduced against total duration. (For details, see Eisler, 1974, 1975.)

This surprising result, that a ratio scale of duration can be constructed from number-free duration reproduction data, should likewise permit the construction of a ratio scale for duration in rats, provided I could train them to reproduce durations.

In Figure 2 the outcomes of duration reproduction experiments with rats and humans can be compared. The figure shows plots for a session of a rat (left) and for a human observer (right). The lower plots are the raw data, reproduced duration plotted against total duration, the upper plots the psychophysical functions. To teach a rat to press the lever when it "experiences" the reproduced duration as equal to the standard, corresponding to the instruction to a human observer "press the button when the second sound has lasted as long as the first," a "correct" region was defined. It was bounded for each standard by the shortest and longest reproduction made by any of the human observers for the standard in question. Responses within this region were reinforced. Naturally, not all ten durations could be presented from the start. The procedure had to proceed in steps, starting with two or three durations without overlap of the correct regions. Originally, I had thought I would have to narrow down this region around the rat's mean response, but this proved unnecessary; the rats seemed to make out the reproduction task quite well.

Communication with rats is even more difficult than with students of psychology. The experiment had to be carried out in three phases. The first was magazine training, the third the duration repro-

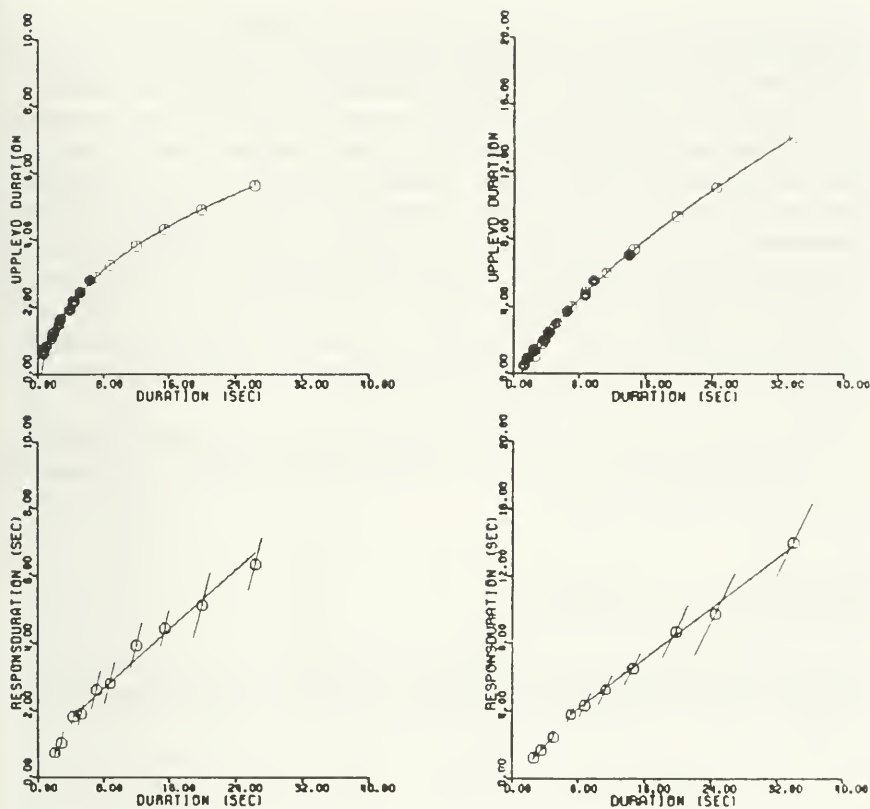


FIGURE 2. Duration reproductions from a rat (*left*) and a human observer (*right*). The *upper panels* describe the psychophysical power function, subjective versus physical duration. Black dots indicate goodness of fit. *Lower panels* show plots of reproduced duration Φ_r versus total duration Φ_t , together with the fitted straight lines. The slanted lines around the points are standard deviations.

duction experiment proper (the result of which is shown in Figure 2) and the second a training phase on which we shall dwell because it offers two examples of serendipitous findings in a narrower sense.

Experimental details can be found in Eisler (1984b, 1984c). Here I shall give only a brief description. We will remember that a trial consisted of an interrupted sound, with a standard duration before and a reproduced duration after the interruption. Trials were separated by a pause of 30 s and ended with access to sweetened milk as reinforcement, contingent upon "correct" behavior of the rat. What the rats had to learn in Phase 2 was to refrain from lever pressing during the standard duration, i.e., during the sound *before* the inter-

ruption, and to press the lever *after* the interruption. The aim of this procedure was to train the rats to attend to the interruption, which in Phase 3 was to be the demarcation between the offset of the standard and the onset of a duration to be terminated by a lever press when experienced as of the same length as the standard. Remember also that all the ten standards were randomly presented in the same session, a session comprising about 50-60 trials. A lever press during the standard gave rise, in addition to forfeiture of the reinforcement, to a prolongation of the sound by 4 s (the "punishment time"), in order to avoid superstitious chaining. Figure 3 shows two cycles (a cycle consists of a trial plus the long pause and possible reinforcement), a correct cycle in the upper panel and one with faulty lever presses in the lower.

The main interest was in the last phase, the duration reproduction phase, see Figure 2, but I was also curious about the course of acquisition during Phase 2.

KNOWING BEFORE DOING

A first observation was that the mean latencies between the end of the interruption and the lever press that terminated the sound (de-

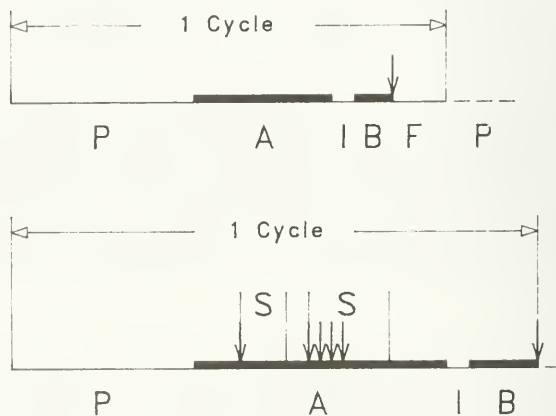


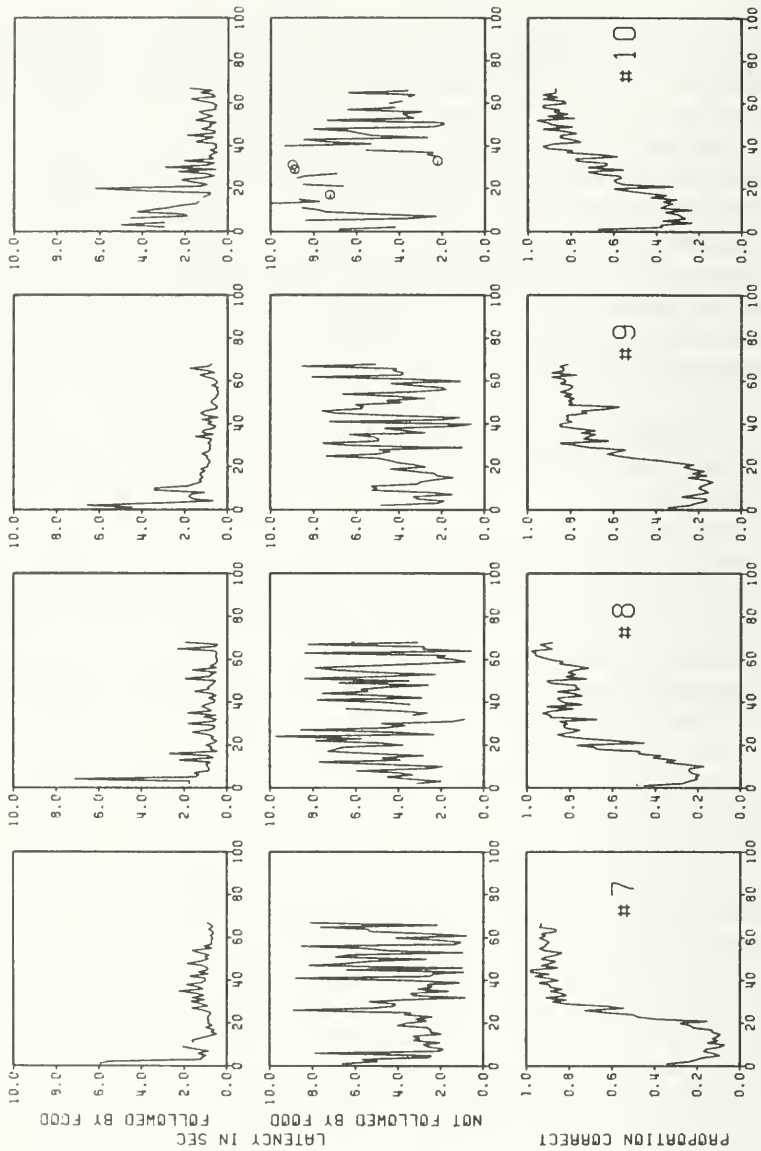
FIGURE 3. Cycle with reinforced (*upper panel*) and nonreinforced (*lower panel*) behavior patterns. The thick lines indicate the ongoing sound, vertical arrows donate lever presses, P the long pause, A the presented standard duration (in the lower panel lengthened by the punishment time S), I the interruption of the sound, B the latency between offset of the interruption and the following lever press, and F the feeding period. (From Eisler, 1987. Copyright 1987 by Elsevier Science Publishers. Reprinted by permission.)

noted by B in Figure 3) decreased with the number of sessions. The data treatment of this experiment was done while I was a Fellow-in-Residence at NIAS (Netherlands Institute for Advanced Study in the Humanities and Social Sciences) at Wassenaar. I happened to mention this finding to Eddy Roskam at the Katholieke Universiteit, Nijmegen, and he suggested separate looks at the latencies for trials with reinforced and not reinforced behavior, respectively. The result was astonishing: the latencies for correct trials (no lever press during the standard; reinforcement after the lever press) were short, almost from the beginning of Phase 2, whereas trials in which the rat had pressed the lever during the standard, so that no reinforcement was forthcoming, showed long (irregular) latencies throughout Phase 2 (See Figure 4). The decrease of the latencies during the course of Phase 2 was thus *not* a *general* decrease but reflected the decreasing frequency of trials with one or more incorrect lever presses, with their long latencies.

Note that the lever press terminating the sound occurred *before* the possible reinforcement, which thus could not have affected the preceding latency. Figure 4 also shows the acquisition curves, which demonstrate that the dissociation between short and long latencies had taken place before the rats' learning not to press the lever during the standard. The conclusion is that the rats knew of their faulty lever presses before they had learned to suppress them: "Knowing before doing" (Eisler, 1984b). This finding was not only unexpected, it was an unintended side effect of the main investigation and thus indeed serendipitous.

I sent a manuscript describing this finding to the *Journal of the Experimental Analysis of Behavior* where, after some linguistic changes (Skinnerian is not my psychological mother tongue), it was accepted, or so I thought. What happened, however, was that the journal acquired a new editor, who proposed a different explanation of the rats' behavior (that longer latencies were found for longer standard durations) and also pointed to a recent article by Shimp (1983) that did not seem to square with my findings. So I had to add an appendix to rule out the editor's hypothesis. Furthermore, I wrote another paper explaining Shimp's results (Eisler, 1984a): Serendipity on a higher level.

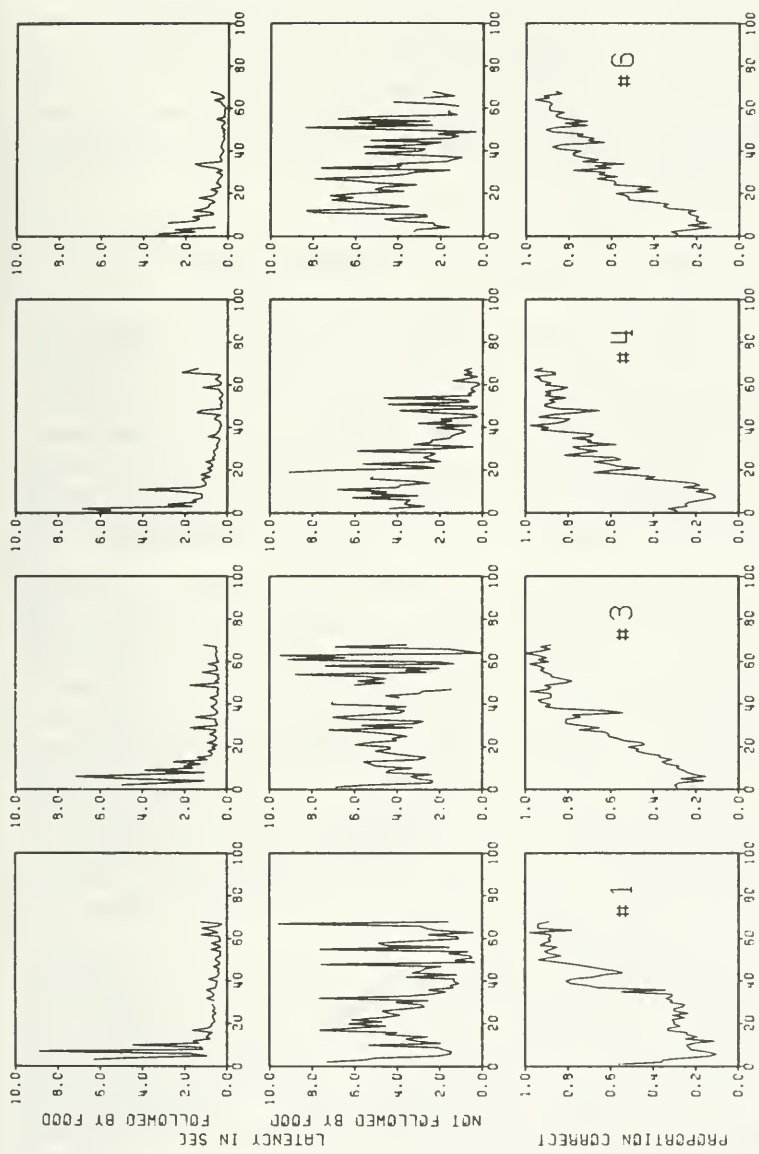
The hypothesis that longer durations entailed longer latencies could be refuted in the following way. As mentioned before, one lever press during the standard (A in the upper part of Figure 3) had the consequence of lengthening the sound before the interruption by 4 s. This made it possible to compare the latencies with and without an erroneous lever press for roughly the same durations A of the sound before the interruption. For instance, the standard duration 4.5 s approximates the 5.3 s which is composed of the standard 1.3 s and the



LATENCY IN SEC

NOT FOLLOWED BY F00D

PROPORTION CORRECT



SESSIONS

FIGURE 4. Acquisition curves (proportion of correct trials per session as a function of session number) and mean latencies with and without impending food as a function of session number for eight rats. Latencies exceeding 10 s are omitted. (From Eisler, 1984b. Copyright 1984 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.)

punishment time of 4 s. Another example is the standard of 20 s and the combination of the standard of 14.8 s and the punishment time 4 s, yielding 18.8 s. An overall scrutiny of the pertinent latency data showed short latencies for trials with no lever press before the interruption, and long latencies for trials with one lever press, independent of the length of duration A.

TEMPORAL HOLES IN THE LATENCY DISTRIBUTIONS

Another problem is the distribution of faulty lever presses, in particular the latency from the start of the standard to the first lever press (before the interruption). For instance: does the rat's impatience grow with the length of the standard, thereby increasing the probability of such a faulty lever press? The equipment could not be used to determine the point in time of a lever press, only whether one (or more) had occurred in any one trial. Thus the distribution had to be inferred.

The longest standard was 20 s. This time interval was divided according to the differences between successive standard durations: 0-1.3 s, 1.3-1.8 s, . . . 14.8-20 s. The frequencies of trials with lever presses for any one such interval was determined as the difference between the frequency of the longer standard (say 1.8 s) and the next shorter one (1.3 s). A scrutiny of the data indicated that the distribution was irregular and that lever presses did not occur during certain of these intervals, namely those in which the frequency of trials with incorrect lever presses was about the same for two consecutive standard durations, e.g., for 1.3 and 1.8 s, so that the difference in frequencies for the interval in question (here the .5 s interval between 1.3 and 1.8 s) was close to zero.

The model describing *random* (first) lever presses is a negative exponential distribution with the cumulative distribution function $1 - e^{-\lambda t}$, where t is the observation period and λ the probability per time unit of the event, here a lever press. The parameter λ decreases with the number of sessions, because the animals learn to suppress lever pressing during the standard, but is assumed to be constant during a session. This model was rejected by a χ^2 test for all rats (fitting an average λ), which was to be expected from the study of the data mentioned above. In order to get a better description of the rats' behavior, the model was changed by replacing the clock intervals t by latent durations T , so that the negative exponential distribution holds. χ^2 was minimized by fitting 9 T values (for the intervals between successive standards) and 68 λ values for each rat. (Phase 2 comprised 68 sessions.) Figure 5 shows the latent duration T plotted against the clock durations t in log-log coordinates. What should be

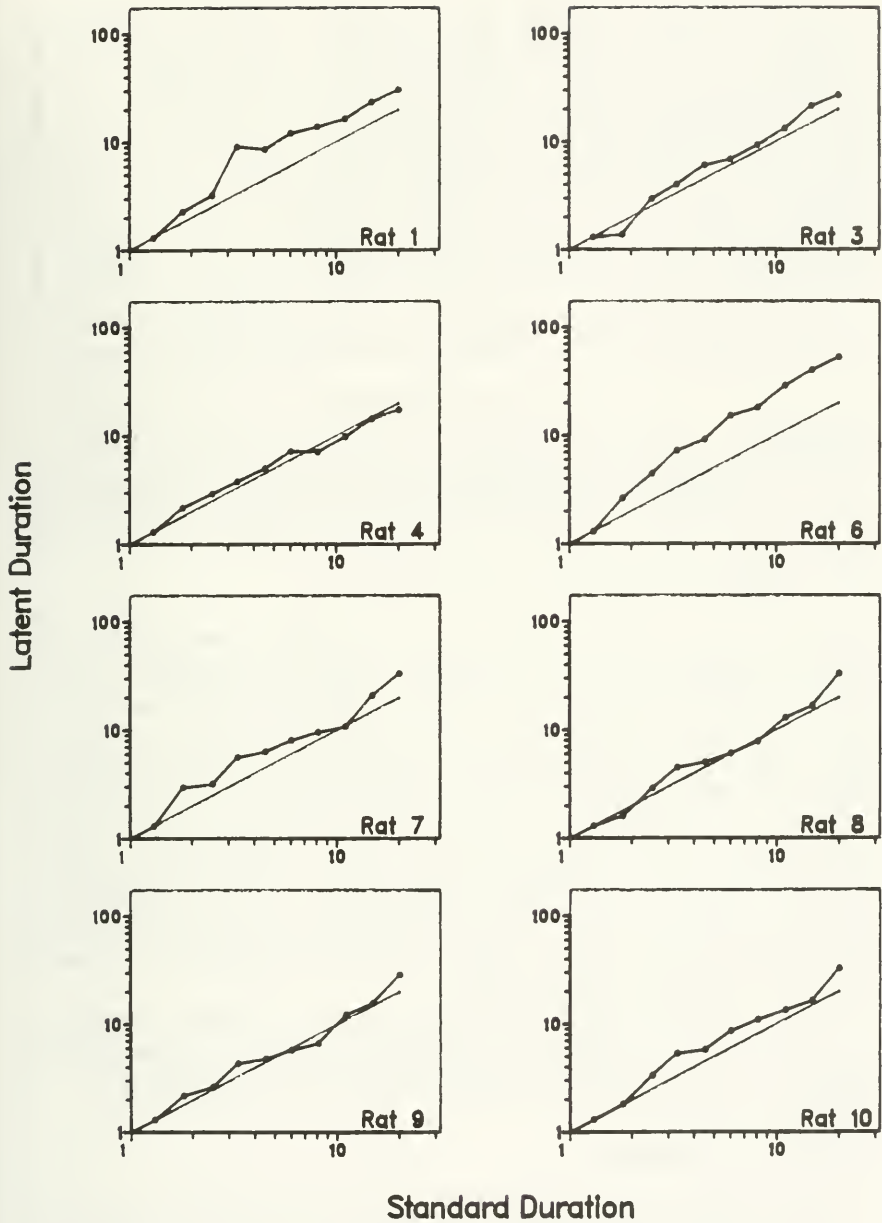


FIGURE 5. Latent durations T plotted against presented durations t for eight rats. The straight line indicates agreement between the two variables. Its slope should be compared with the slopes between adjacent points. (From Eisler, 1987. Copyright 1987 by Elsevier Science Publishers. Reprinted by permission.)

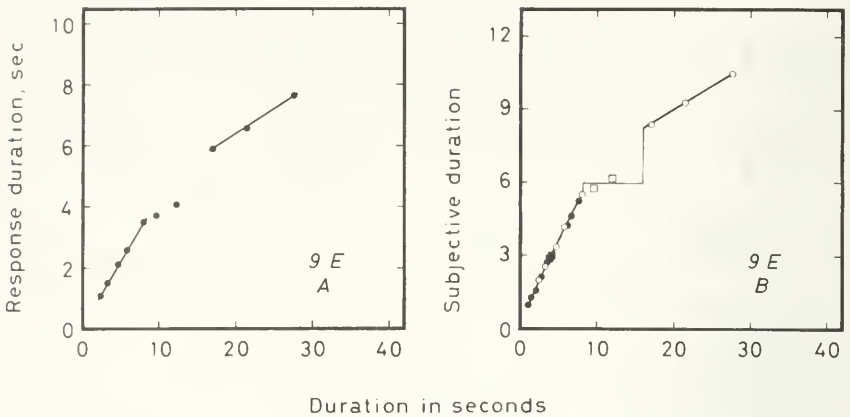


FIGURE 6. Duration reproduction for a human observer. Right panel: The psychophysical power function. Left panel: Reproduced duration vs. the sum of the standard and reproduced durations. (From Eisler, 1987. Data from Eisler, 1975. Copyright 1987 by Elsevier Science Publishers. Reprinted by permission.)

observed are the slopes. The lower the slope for a given interval, the shorter the latent duration compared to the clock duration. We see that for 7 of the 8 rats there are intervals with a slope close to zero. (For 5 rats this was the case for the interval 3.3-4.5 s.) In some sense these intervals do not exist for the rats; no lever presses took place. They are “temporal holes” in the latency distributions. This finding may be connected with the breaks and similar “holes” found in experiments with human observers, see Figure 6. An explanation could be a switch from one neural loop, corresponding to a certain duration, to the next (Thatcher, 1979; Thatcher & John, 1977). Details of this study can be found in Eisler (1987). Again, the finding of “temporal holes” is an instance of serendipity.

To conclude: it pays to keep your eyes open for unexpected findings in your experiments! Even if the problem focuses on steady-state behavior, a study of the acquisition phase may be worthwhile.

ACKNOWLEDGEMENT

This work was supported by the Swedish Council for Research in the Humanities and Social Sciences.

REFERENCES

- Eisler, H. (1974). The derivation of Stevens' psychophysical power law. In H. R. Moskowitz, B. Scharf, & J. C. Stevens (Eds.), *Sensation and measurement* (pp. 61-64). Dordrecht, Holland: Reidel.

- Eisler, H. (1975). Subjective duration and psychophysics. *Psychological Review*, *82*, 429-450.
- Eisler, H. (1984a). Comments on Shimp's double dissociation between knowledge and tacit knowledge. *Journal of the Experimental Analysis of Behavior*, *41*, 341-344.
- Eisler, H. (1984b). Knowing before doing: Discrimination by rats of a brief interruption of a tone. *Journal of the Experimental Analysis of Behavior*, *41*, 329-340.
- Eisler, H. (1984c). Subjective duration in rats: The psychophysical function. In J. Gibbon & L. Allan (Eds.), *Timing and time perception: Annals of the New York Academy of Sciences*, Vol. 423 (pp.43-51). New York: New York Academy of Sciences.
- Eisler, H. (1987). Timing of rats' lever pressing when learning not to press: Temporal holes in the latency distributions. In E. E. Roskam & R. Suck (Eds.), *Progress in mathematical psychology- 1* (pp. 219-232). New York: Elsevier Science Publishers B. V. (North-Holland).
- Shimp, C. P. (1983). The local organization of behavior: Dissociations between a pigeon's behavior and self-reports of the behavior. *Journal of the Experimental Analysis of Behavior*, *39*, 61-68.
- Thatcher, R. W. (1979). The neural representation of experience as time. In G. G. Haydn (Ed.), *Experience forms: Their cultural and individual place and function* (pp. 3-22). The Hague: Mouton.
- Thatcher, R. W., & John, E. R. (1977). The neural representation of time. *Foundations of cognitive processes* (Vol. 1, pp. 165-179). Hillsdale, NJ: Erlbaum.