心理学报 2010, Vol. 42, No.1, 138−158
Acta Psychologica Sinica DOI: 10.3724/SP.J.1041.2010.00138

Screening for Learning and Memory Mutations: A New Approach

C. R. Gallistel[1], A. P. King[2], A. M. Daniel[1], D. Freestone[1], E. B. Papachristos[1], F. Balci[1], A. Kheifets[1], J. Zhang[1,3] , X. Su[1], G. Schiff[1], H. Kourtev[1]

([1] Department of Psychology & Center for Cognitive Science, Rutgers University, New Brunswick, NJ, USA)

([2] Department of Mathematics and Computer Science, Fairfield University, Fairfield, CT, USA)

([3] Behavioral Genetics Centre, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, P.R. China)

Abstract: We describe a fully automated, live-in 24/7 test environment, with experimental protocols that measure the accuracy and precision with which mice match the ratio of their expected visit durations to the ratio of the incomes obtained from two hoppers, the progress of instrumental and classical conditioning (trials-to-acquisition), the accuracy and precision of interval timing, the effect of relative probability on the choice of a timed departure target, and the accuracy and precision of memory for the times of day at which food is available. The system is compact; it obviates the handling of the mice during testing; it requires negligible amounts of experimenter/technician time; and it delivers clear and extensive results from 3 protocols within a total of 7-9 days after the mice are placed in the test environment. Only a single 24-hour period is required for the completion of first protocol (the matching protocol), which is strong test of temporal and spatial estimation and memory mechanisms. Thus, the system permits the extensive screening of many mice in a short period of time and in limited space. The software is publicly available.

Key words: temporal memory; matching; phenotyping; mutation; automation

In the 1970's, Seymour Benzer pioneered the use of genetic methods to discover the molecular basis for mechanisms that perform critical behavioral functions. He and his students used what is now called *forward genetics* in *Drosophila melanogaster* to discover the molecular basis for the circadian timing of behavior and to search for the molecular mechanism of learning and memory (Weiner, 1999).

The general idea behind a forward genetics approach, as it applies to behavioral mechanisms, is to use a behavioral test to screen for heritable malfunction in a specified mechanism. When a heritable malfunction is found, classical genetic methods are used to locate the mutated gene on a small segment of a chromosome. Then, molecular genetic methods are used to identify and sequence the gene. Knowledge of this sequence enables molecular biologists to fashion an ever-expanding set of molecular tools that may be used to elucidate the molecular and cellular biology of the mechanism within which the protein coded for by the sequenced gene functions. The discovery of one important gene usually leads to the discovery of other genes in the set of genes that code for different parts of the mechanism or control its assembly. Thus, the critical first step—the behavioral screen that identifies a mutation in a gene that codes for a critical component of the mechanism—gives biologists the end of a thread that they may follow down into the workings of the molecular mechanism and from that mechanism up into the cellular and system's level mechanisms.

The Benzer-originated work on the circadian clock met with quick success, giving rise to a steadily growing understanding of the clock mechanism at the molecular and cellular level (Antle & Silver, 2005; Maywood et al., 2007; Takahashi, 2004). The genetic approach to the

Received date: 2009-05-05

Correspondence should be addressed to Dr. C.R. Gallistel, Department of Psychology & Center for Cognitive Science, Rutgers University, New Brunswick, NJ, USA, 08854-8020. E-mail: galliste@ruccs.rutgers.edu Fax: (1)732-445-6715

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

mechanism of learning and memory has given rise to a large literature (Dubnau, Chiang, & Tully, 2003; Kandel, 2004), but it is not clear that it has led us to the molecular mechanism of memory (Eichenbaum, 1996; Koch, 1999; Leil, Ossadtchi, Cortes, Leahy, & Smith, 2002; Mercer et al., 2008; Shors & Matzel, 1997). It is instructive to consider why the first effort has been more unequivocally successful than the second.

In the first case, there was a coherent conception of the sought-for mechanism (Bruce & Pittendrigh, 1956; C.S. Pittendrigh, 1960; C. S. Pittendrigh, 1965; Richter, 1965): it was an oscillatory process whose period was to a surprising degree independent of temperature and other characteristics of the environment, including the period of the light-dark cycle. Its phase could be altered by time-giving signals from the light-dark cycle, but not its period. The sign and magnitude of the behaviorally measured phase-shift in response to a signal from the light-dark cycle depended on the phase of the activity cycle when the signal arrived (the phase-response function). The period of the free-running cycle and its phase-response function can be measured at the molecular and cellular level as well as at the behavioral level, which makes it possible to check whether a given molecular or cellular mechanism has properties in quantitative accord with the properties revealed by behavioral measurements. Quantitative correspondence between behavioral and cellular or molecular measures is powerful evidence in establishing the identity between a molecular or cellular mechanism and a mechanism revealed through its behavioral effects (C.R. Gallistel, Shizgal, & Yeomans, 1981).

What is Memory?

In the case of the memory mechanism, the guiding conception of the sought-for mechanism is incoherent. On the one hand, researchers have been looking for a symbolic memory, a mechanism that encodes and preserves in retrievable form information revealed by experience. On the other hand, researchers have been looking for the mechanism of association formation. In looking for the associative mechanism for information encoding, researchers are seemingly oblivious to the fact that the associative theory of memory is anti-representational; it is an alternative to the hypothesis that there is a symbolic memory mechanism, a mechanism that encodes information and carries it forward in time in a readable form. The marriage of these contradictory assumptions is seen in the following recent quote: “Memories are encoded by a specific pattern of activity that is unique to the information being processed and stored. Memory formation is almost certainly achieved at the synaptic junctions between neurons through the process of long-term potentiation (LTP), whereby synaptic communication between two simultaneously active neurons becomes stronger.” (Thompson & Mattison, 2009, in *Nature*, p. 296)

The just-quoted sentences reflect mainstream thinking about the mechanism of memory. They assert both that the memory mechanism encodes information and that it is an associative process, wherein the connections between simultaneously active neurons are made stronger. The implication is that these two assertions are connected by an intelligible hypothesis specifying how alterations in the connections between neurons may encode information in a readable (retrievable) form. But there is no such hypothesis—and that for good reason: The hypothesis that memory is the alteration in the connectivity between neurons consequent upon the temporal coincidence of experienced events is a translation into neurobiological terms of the associative theory of learning and memory. That theory has always stood in opposition to the idea that the brain forms a symbolic representation of the experienced world—the hypothesis that the brain “stores information.” Associative connections are not suited by either their form or their causation to fulfill the role of information-encoding symbols, the role that stored bytes play in a computer and that nucleotide sequences play in the conveyance of inherited information (C.R. Gallistel, 2008; C.R. Gallistel & King, 2009). Associative bonds, that is,

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

signal conducting connections forged by experience, are not a medium that can encode information in an accessible form.

The incoherence in the conception of what researchers are looking for goes hand in hand with the absence of behaviorally measured quantitative properties of the learning and memory mechanism. As indicated in the already quoted sentences, it has always been taken for granted that the temporal pairing of signals is essential to the associative learning mechanism: "Associative learning by definition depends on temporal pairing between stimulus and reinforcement."-- (Quinn, 2005, in *Nature Neuroscience*, p. 1639). One would think, therefore, that there were behavioral measurements defining the "window of associability," measurements specifying what constitutes temporal pairing. Surprisingly, the window of associability, the critical interval, has never been convincingly measured — for any subject species in any learning task — not even, for example, for the rabbit in the widely used eyeblink conditioning task (Balsam & Gallistel, 2009; C.R. Gallistel, 2007; C.R. Gallistel & Gibbon, 2000; Rescorla, 1972). Thus, there are no behavioral measurements against which to compare the results from measurements on the critical role that temporal pairing measured in milliseconds appears to play in LTP and LTD (Markram, Lübke, Frotscher, & Sakmann, 1997). There is no remotely comparable behavioral finding.

It is unclear even how to define temporal pairing in some behavioral paradigms that are widely used to investigate the neural basis of learning and memory. In the water maze, for example (R. G. M. Morris, 2003; Wood et al., 2005; Zhang, Zou, He, Gage, & Evans, 2008), the animal learns where the submerged platform is. There is little work analyzing what exactly the animal learns when it learns "where the platform is." However, it seems reasonable to suppose that it learns how far the platform is from the wall of the circular tank and in what angular direction from the center of the tank relative to the enclosing space (the room in which the tank is located), because direction and distance are fundamental to the representation of spatial relations (C.R. Gallistel, 1990). What is temporally paired with what when an animal abstracts from its experience and commits to memory the direction and distance of a location is a question that has not been addressed. If temporal pairing is in fact what drives the associative mechanism and if the associative mechanism is the mechanism of spatial memory, then the question appears unavoidable. Also unavoidable is the question of how changes in synaptic conductances might encode remembered directions and distance (that is, vector variables, see C.R. Gallistel & King, 2009, for an extensive discussion of this question).

Because the associative theory of learning focuses on presumed changes in the strengths of the connections, another physiologically meaningful variable that one might try to measure is the rate of change in connection strength. In the behavioral neurogenetics literature, the rate of learning in mutant and wild-type strains has often been estimated by plotting a group average measure of performance—for example, the mean latency to find the submerged platform—as a function of trials.

A problem with the rate-of-learning measure is that the gradual approach to asymptote seen in group-average plots is an artifact of averaging across trials and subjects (Papachristos & Gallistel, 2006). In the individual subject, the transition from the initially low and/or slow performance to later asymptotic performance usually occurs in a single step (C.R. Gallistel, Balsam, & Fairhurst, 2004; R. W. Morris & Bouton, 2006). The trial on which the step occurs and the size of the step vary greatly between subjects, even in highly inbred strains of mice (Papachristos & Gallistel, 2006). Step learning "curves" for the individual subjects are seen in most of the basic animal learning paradigms, including paradigms as diverse as the rabbit eyeblink, rodent and pigeon Pavlovian conditioning (C.R. Gallistel, Balsam, & Fairhurst, 2004; R. W. Morris & Bouton, 2006; Papachristos & Gallistel, 2006), the rodent water maze (C.R. Gallistel, Balsam, & Fairhurst,

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

2004), and the rodent learning of the peak procedure (Balci et al., 2009). Averaging across the steps produces the gradual curves that are often used to compare rates of learning in groups of subjects given different neurobiological manipulations. Because the form of these learning curves is an artifact of averaging across subjects, rate parameters extracted from them do not reflect meaningful quantities within individual subjects. Whatever the physical changes that mediate memory are, they occur in the brains of individual subjects, not in some insubstantial memory "ether" common to a random sample of mice.

What is Learning?

A conceptual problem that is to some extent implicit in what has already been said is that the associative theory treats "the" learning mechanism and the memory mechanism as one and the same. In an associative theory, learning is the process of association formation by virtue of the temporal pairing of the neural signals generated by events. Memory is the resulting connection. This contrasts with theories in which memory is the mechanism by which the information extracted from experience is carried forward in time in a computationally accessible form, so that it may inform subsequent behavior in the indefinite future. In such theories, the learning mechanisms and the memory mechanism(s) are utterly different mechanisms (C.R. Gallistel, 2008; C.R. Gallistel & King, 2009). There might be a universal memory mechanism, because this mechanism performs the same simple function in every domain in which it operates. It carries information from one location in time to later locations in time. Its function is closely analogous to that of the action potential, whose function is to carry information from one location in the nervous system to another. If a universal mechanism—the action potential—serves to convey any kind of information from one location to another within the nervous system, why should we not imagine a universal mechanism for conveying any kind of information from one time to a later time? Generally speaking, mechanisms for conveying information are indifferent to content (what it encoded). When it comes to the transmission and storage of information, it's all bits.

On the other hand, the learning mechanisms in this latter kind of theory must be specific to the sort of thing that was to be learned and the nature of the data from which it must be induced (C.R. Gallistel, 1999b). Learning mechanisms operate differently on information from different kinds of data. The dead reckoning computations by which animals keep track of their location is very different from the parameter-setting computations by which they learn the solar ephemeris, and both of these are very different from the computations that enable them to solve the nonstationary, multivariate time series problems posed by Pavlovian conditioning paradigms (Gallistel, 1999b, 2002, 2003; Gallistel & King, 2009). Even when the same kind of information is extracted, for example, the direction of a distal stimulus, we know that very different mechanisms come into play depending on whether the proximal stimulus is auditory or visual (C.R. Gallistel, 1999a).

If learning is a purely associative process, then it is not symbolic in nature and therefore cannot be understood in information-theoretic terms. In associative theories, memory has not symbolized content. The information that an experience communicates to a receiver is measured by the reduction in the receiver's uncertainty regarding some state of the world (Shannon, 1948). A purely associative receiver knows nothing of uncertainty. Because it does not have symbols to encode either states of the world or probabilities, a fortiori, it does not have symbols that specify probability distributions over possible states of the world. When a receiver lacks the wherewithal to specify a probability distribution, then it has no measurable uncertainty (no symbolic entropy). A fortiori, there is no way of changing its uncertainty.

On the other hand, if learning is the extraction from experience of behaviorally useful information, then it does not make sense to assume there is a

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

single computation capable of extracting any kind of information from any kind of sense data. Under this latter assumption, learning mechanisms must be tailored to the problems that they solve, just like the mechanisms that we find everywhere else in organic structure. In sum, in non-associative theories of learning and memory, learning mechanisms must be domain specific, but there is no reason why the memory mechanism should not be universal.

The Handling Problem

Turning from the conceptual problems that beset current attempts to use behavioral neurogenetic screening to the practical problems, the methods in current use do not lend themselves to large-scale high throughput screening. Most of them require the handling of the subjects in the course of the training and testing. This is doubly undesirable. It consumes large amounts of experimenter and technician time in the obtaining of small amounts of data. And, it seriously stresses the subjects. Most strains of mice react badly to handling, although the extent, duration and manifestations of handling stress vary greatly between strains. Moreover, the skill with which the mice are handled varies greatly between laboratories and even between personnel within laboratories. Reactions to having been handled and the anticipation of soon being handled again may take a long time to subside once a mouse has been placed in a test environment. These reactions to handling interfere with and contaminate almost every kind of behavioral measurement and observation.

Our Method

The just reviewed theoretical and practical considerations have led us to our method of screening mutant mice strains for malfunctions in basic mechanisms of cognition, with particular emphasis on the mechanism of memory, which we assume to be as central to a brain's computational capacity as DNA is to life (C. R. Gallistel & King, 2009).

- We target mechanisms, like the circadian clock, for which one can make physiologically meaningful behavioral measurements. Behavioral measures are physiologically meaningful when one can reasonably imagine making comparable measurements at the systems, cellular and molecular levels of analysis. Measuring the free running period of the circadian clock is an example. Behavioral determinations of spectral sensitivity curves in vision (e.g. Foster, 1993) and of whole-nerve conduction velocities and refractory periods are other examples (C.R. Gallistel, Shizgal, & Yeomans, 1981). Kelvin famously observed that when you cannot measure something you have very little understanding of it. Our version of this is that if you cannot make physiologically meaningful measurements from some behavioral phenomenon, then you have little hope of finding its mechanism,
- We want highly automated procedures that eliminate handling of the mice during the period when behavioral measurements are made and give as many measurements as possible in as little time as possible.
- Because we believe that different learning mechanisms are likely to make use of a common molecular mechanism for carrying information forward in time in computationally accessible form, we need to develop screens for several different kinds of simple quantitative learning. A malfunction in a memory mechanism common to them may reveal itself in a behaviorally measurable quantitative aberration that is common to them all.

Targets

Our research targets the interval timing mechanism, whose behavioral investigation was pioneered by Gibbon and Church (Church, 1984; Gibbon, Church, & Meck, 1984) and the mechanisms for estimating probabilities (relative frequencies) and the proportions obtaining between them. The physiologically meaningful quantities that we measure are the accuracy and precision of the individual subject's representation of these

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

objective quantities (duration and relative frequency and proportion). We have developed paradigms for measuring these quantities rapidly in a live-in environment, which eliminates the handling of the mice (Figure 1). By automating every aspect of the situation, including the data analysis, which is conducted in quasi real time, we make it possible to do large scale screening with an equipment investment no larger than is required for many major molecular and neurobiological experimental programs.

We use the matching paradigm to measure the accuracy with which the mouse estimates the average intervals between randomly scheduled pellet releases into two different hoppers and the accuracy with which it represents the proportion between these average intervals. In the matching paradigm, the mouse adjusts the expected durations of its visits to the two hoppers so that their ratio (the proportion between the two expectations, which we call the temporal investment ratio) matches the ratio of the rate of pellet release (the income ratio). Mice reliably exhibit matching within the first few hours in a new test environment, a period during which they may remain so wary of the new environment (and perhaps so stressed by the handling required to put them in it) that they eat only a few of the pellets they obtain by poking into the feeding hoppers (C.R. Gallistel et al., 2007).

We use the "switch" paradigm (Balci et al., 2008) to measure the accuracy and precision with which the mouse represents durations and the accuracy with which it represents a probability (relative frequency). In this paradigm, a trial begins with the illumination of a trial-initiation hopper. When the mouse pokes into this illuminated hopper, its light goes out and the lights come on in the feeding hoppers that flank it (see Figure 1). With some relative frequency, the trial terminates with the delivery of a pellet to, say, the left hopper after a fixed delay of, say, 3 s. With the complementary relative frequency, it terminates with the delivery of a pellet to the other hopper, after a longer, fixed delay (say, 9 s). The mice soon learn to poke first into the short-delay hopper and to switch to the long-delay hopper on those trials (long trials) when the short delay expires without the release of a pellet. Measures of the accuracy and precision of interval timing and interval memory are simply derived from the cumulative distribution of switch latencies on these long trials. This distribution shifts toward or away from the short delay according as it is less or more probable (Balci, Freestone, & Gallistel, 2009).

By scheduling foraging sessions of a few hours each at different times of the day and night, we are also able to measure subject's memory for the circadian phase (the time on their internal clock) at which these sessions begin. The subject's visits to the (inactive) hoppers pick up noticeably before the time when the hoppers become active.

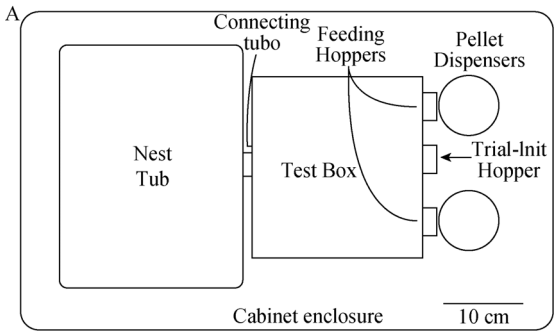

Figure 1. A. *Plan of live-in test environment. A nest tub communicates with a Med Associates™ Mouse Test Box by way of a connecting tube. Test box has three illuminable hoppers monitored by infrared beams. The lateral hoppers, the "feeding hoppers", are connected to pellet dispensers. In some protocols, the illumination of the middle hopper, the "trial-initiation hopper," signals that the mouse may initiate a trial by poking into that hopper. The first poke into the illuminated trial-initiation hopper extinguishes the illumination in that hopper and illuminates one or both of the flanking feeding hoppers.*

Data Analysis

These automated protocols operating in a live-in environment generate large amounts of data. We record all of the stimulus events and the times at which they occur and all of the interruptions of the infrared beams at the hopper entrances and at the two ends of the tube that communicates between the nest tub and the test box and the times at which these beam interruptions occur. There is a

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

substantial risk of drowning in the data. A systematic, well thought out approach to archiving these raw data files and the many, often quite elaborate, analyses of them is essential. To this end, we have created an open source Matlab™ toolbox for the analysis of extensive time-stamped event records. This open-source toolbox may be downloaded from http://cognitivegenetic.rutgers.edu/tslib/download.html, where a link to extensive tutorial material will also be found. Included in that downloadable material is the code that produced the results we here report, including the graphics code.

The principles that guided the development of the toolbox are:

- *Keep it all together*: The raw data and the results extracted from them should be processed in such a way that they are inseparable one from the other and from the code that governed the operation of the testing equipment and the logging of the data when the data were gathered. This latter process-control code is an essential part of the experimental protocol.
- *Make a clear trail*: It must be possible to regenerate the published analyses from the raw data, that is, for the trail from the raw data to the figures and tables and numerical values that appear in the published reports to be retraced without difficulty.
- *Make complex analyses easy.* One can pose many different questions when one has a rich time-stamped database, discovering important results that were not foreseen when the experiment was designed. However, this is likely to happen only when it is relatively easy for an appropriately trained researcher to pose an unforeseen question to the data and quickly get an answer. If one has to write more than a few lines of code to get an answer to a new question, there will be many fewer questions posed to the data.
- *Better buy than build*. The analytic software should be embedded in a powerful general purpose, widely used, securely and extensively supported programming, statistics, and graphing system, such as Matlab™ or Mathematica™ or the "R" system.

In conformity to the last principle, we decided to have our data-analysis software take the form of a custom Matlab™ toolbox. Matlab™ is a proprietary programming platform, but our toolbox is non-proprietary, open source Matlab code.

In conformity to our first principle—keep it all together—the software system puts the raw data and all the results that come from analyses of that raw data into a single "structure." A structure in Matlab is a data type that provides flexibility and intelligibility in organizing vast and diverse data structures, while making all data, both numerical and textual, accessible to computation. It is a hierarchically structured set of data fields with user-chosen names. The flexibility comes from the hierarchical structuring and from the fact that there are no restrictions on what can be put in a field; a field may contain anything from a single number to a long text to an another complex structure (that is, structures may be embedded within other structures). The intelligibility comes from the hierarchical arrangement of numerically indexible fields. The user creates the field names, just as they would create headings in a spreadsheet, with the advantage that fields in a Matlab structure are numerically indexible [Subject(3), Trial(21), etc].

The data are accessed by way of the hierarchically structured field names. For example, the command:

mean(Experiment.Subject(3).Session(5).VisitDurations)

computes the mean of the column (that is, field) of visit durations during the fifth session for the third subject. "Experiment" is the name of the entire structure. "Subject" is a field with a different index number for each subject. Under each instance of the "Subject" field, there are indexed "Session" fields. Under each of these, there can be a large structure with many additional layers of fields. In this simple example, a field called 'VisitDurations', which is immediately surbordinate to the Session field,

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

contains the durations of successive visits to some location of interest, such as a feeding hopper. As this command illustrates, these data are accessed by way of the field hierarchy.

A Matlab™ structure is analogous to the hierarchically organized data arrays that users create in spreadsheets. The headings above the columns of numbers in a spreadsheet tell you what the numbers in a column represent and the hierarchical arrangement of sub-headings tells you how the columns of numbers are related. Spreadsheet treatment of voluminous time-stamped data is not feasible. There are too many columns; they are too long; the hierarchical structure is too complex; the computations provided by the spreadsheet are not diverse and powerful enough; those provided take much too long when they operate on really large masses of data; and the graphic presentation resources are not sufficient for scientific purposes.

The numbers specifying visit durations in the above example are not in the raw data. Like most numbers of interest, they must be computed from the raw data. These computations begin by finding sequences of (generally non-contiguous) events. Whenever a sought-for sequence is found a statistic, such as the duration of an event, is computed from the time stamps associated with the micro-events that compose the sequence. The statistic is stored within the same structure that contains the data from which it was computed—in a field created by the user. Suppose, for example, that one wants to know the durations of successive pokes into a feeding hopper. The onset of each poke is indicated by a time-stamped beam interruption. The offset of the poke is indicated by the next occurring beam-completion event (when the mouse withdraws its head from the hopper). This latter event will often not be the next event in the sequence of recorded events. Other events, such as pellet deliveries or light offsets or onsets may intervene. To compute the duration of a poke, the data-analysis program must find the onset, find the subsequent offset, then subtract the time stamp of the former from the time stamp of the latter. It must do this for each of the many hundreds of pokes that typically occur in the course of 24 hours.

Behavioral events have hierarchical structure. A sequence of pokes into the same hopper uninterrupted by any events that happened elsewhere (for example, a poke into another hopper or an exit from the test box) constitutes a visit event. The duration of the visit encompasses the durations of all the pokes and interpoke intervals that comprise that visit. High-level data- analysis commands should make it easy to abstract multiple levels of structure from the sequence of time-stamped micro-events, and compute statistics for each level of behavioral structure. Our system accomplishes this by allowing users to define “trials” in extremely versatile ways and then to look for substructure within those trials.

Our concept of a trial was inspired by the practice common in learning experiments of organizing events into trials. However, we have generalized the notion to any sequence of events (and non-events) or even any of several disjoint such sequences. Thus, in our system, “trial” simply means ”stretch of data of interest to the researcher.” To a good approximation, any sequence of non-contiguous events and non-events that a human record scorer could be instructed to look for can define a trial and become the basis for the automated parsing of the raw data into “trials.” For example, a visit to Feeding Hopper 1 might define a trial. It might be defined by the logical OR of the following two sequences:

[PokeOn1 PokeOff1 –PokeOn1 PokeOn2] *or* [PokeOn1 PokeOff1 –PokeOn1 Tube1]

The first sequence detects the mouse coming to Hopper 1, making a sequence of pokes, and then going to Hopper 2, while the second detects the mouse coming to Hopper 1, making a sequence of pokes, and then leaving the test box. Either sequence constitutes a visit to Hopper 1. Figure 2 illustrates the progress of the computation that detects the first of these two constitutive sequences.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

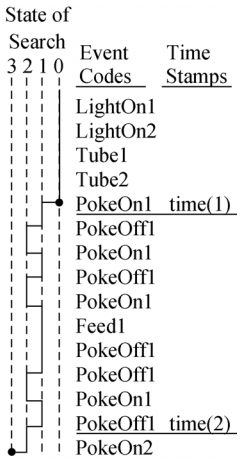

Figure 2. *The progress of the computation that detects one of the two different sequences that constitute a visit to Hopper 1 (see text for the two sequences) is illustrated by the steps in the solid line to the left of a hypothetical sequence of events. When the first event in the definition is encountered, the search steps from the 0 state to the 1 state; when the second is encountered, it steps to the 2 state. When a negating event (in this case, PokeOn1) is encountered, it steps back to the previous state. When the third (positive) event is detected while the search is in State 2, it steps to State 3, which, in this example, is the terminal state. Attainment of a terminal state signifies the discovery of a stretch of data constituting a user-specified "trial" of a particular kind. The user may specify any number of different kinds of trials. Any one kind may be defined by the ORing of any number of different trial-defining sequences. A trial-defining sequence may have several negating events in immediate sequence. Encountering any one of them will step the search back to its preceding state. Once that backward step has been made, further encounters with negating events in that same sequence of negating events will not step the search back to a still earlier state. All of the (positive) events have time stamps, but we indicate here only the two that may be used to compute the duration of the visit.*

The sequence of trials of any given kind constitute an indexed array of fields. Suppose for example, that we have defined two kinds of trials, Hopper1_Visits and Hopper2_Visits. Suppose further that there were 48 visits to Hopper 1 and 27 visits to Hopper 2 in a given session, say, Session(3), by a given subject, say, Subject(8). Then, subordinate to "Experiment.Subject(8).Session(3).TrialHopper1_Visits," one would find:

Trial(1)
Trial(2)
•
•
•
Trial(48)

and subordinate to "Experiment.Subject(8).Session(3).TrialHopper2_Visits," one would find:

Trial(1)
Trial(2)
•
•
•
Trial(27)

Each such sub-field would itself contain a structure of user-specified fields. These would contain user-specified statistics computed from the sequence of the events falling between the initial event and the final event in a trial-defining sequence. The sequence includes all of the events in that stretch of data, not just the events used to define the trial, that is, to pick out that stretch of data. In conformity with our third principle—*make complex analyses easy*—the toolbox allows users to compute almost any computable statistic with one or, at most, two commands. For example, one might use such a command to compute a "VisitDuration" statistic by subtracting the time stamp associated with the first event in the trial-defining sequence—"time(1)" in Figure 2—from the time stamp associated with the second (positive) event in that sequence—"time(2)" in Figure 2. For further example, a single command may create a vector giving the durations of the poke and interpoke intervals within a trial.

Other commands allow the user to combine the statistics from one or more fields at the trial level into fields at the Session level and the statistics from one or more fields at the session level of the structure into fields at the subject level, and statistics from one or more fields at the subject level into fields at the Experiment level.

All of these commands permit the user to specify in powerful and flexible ways, the statistics that are to be computed, using the immense resources of the Matlab platform — in conformity with our *better buy than build* principle. They

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

operate on data contained in previously created fields within the global, all-encompassing "Experiment" structure. They store the results in user-created field within the same structure — in conformity to our *keep it all together* principle. Each trial-oriented command operates on every trial of a specified kind in every session for every subject—unless the user restricts the range of application. The ability to restrict the range of application is also powerful, flexible, and general.

The commands themselves are typically grouped into small sets, using the "cell" feature of the Matlab script editor. This feature allows the user to issue subsets of commands ("cells") without leaving the script editor, while retaining the capacity to run the entire sequence of commands by calling the script itself. A single script file contains the code for the entire data analysis. It is the only file other than the file that contains the Experiment structure. Calling the script from the Matlab command window executes the entire analysis, creating de novo the Experiment structure, filling it with the raw data (assuming that those files are still accessible), creating and filling all the fields within the Experiment structure that contain statistics derived from the raw data. This conforms to our *keep it all together* and our *keep a clear trail* principles. If the archived raw data files have themselves been lost — it has been known to happen! — there is no problem, because the raw data are copied into the Experiment structure prior to any analysis of them. Thus, the script file enables the user or anyone else to reconstruct the analysis that led to the published results, starting either from archived raw data files or from an Experiment structure that contains copies of the raw data. The organization of the data into a single hierarchical structure together with a single script containing the code that generates the structure facilitates the deposition of the results of phenotyping screen into the large on-line data bases where the results of mouse phenotyping are made publicly accessible.

Finally, our toolbox contains a few powerful graphics commands that enable the user to create graphics such as the raster plot in Figure 3, which we have found particularly useful for visualizing behavior while sticking close to the raw data. These supplement (and make use of) the immense graphical resources provided by the Matlab platform.

Automation

Forward genetics requires screening of many different strains. It puts a premium on devising screening systems that are maximally automated, minimizing the amount of experimenter and technician time that must be invested in the screening effort. Our system uses three kinds of software to produce nearly complete automation. The first and third kinds of software have already been described. First, there is the software that implements the behavioral testing protocol by controlling the live-in experimental environment with its standard commercially obtained mouse-testing chamber. This software comes with the commercial obtained test equipment (*better buy than build*). Third, there is our Matlab toolbox, which makes it (relatively) easy to write the powerful data analysis code required to digest the voluminous data. The second kind of software is a shell that bridges between the first and third kinds (Figure 4). It is written in a general-purpose object-oriented scripting language (Ruby), running on a server. The shell looks periodically—how often is specified by the user—at the data file to which the testing software is writing the data. It checks the data for error codes. If it finds them, it sends email messages to a specified list, alerting them to reported errors. It may also be instructed to send an email alert when no data have been written for some suspiciously long interval (usually several hours), as this also makes it likely that there is a problem in the operation of the test apparatus (caused by for example a power failure or a malfunction in the computer controlling the test apparatus).

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

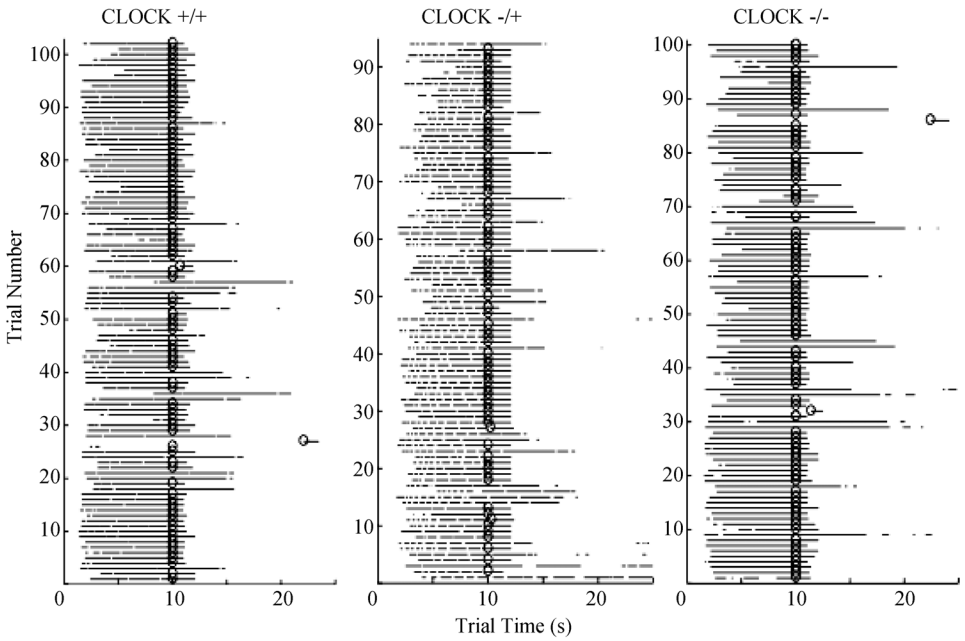

Figure 3. The command TSraster in our Matlab toolbox generates raster plots like these. Each plot shows the performance of a mouse over approximately 100 trials in a "peak procedure" protocol. In this protocol, the illumination of the hopper signals that a pellet will be delivered in response to any beam interruption at or after 10 seconds has elapsed (FI 10 s). On some trials, called probe trials, the pellet is not delivered; the hopper illumination persists for 3 to 4 times the delivery latency, but mice learn to stop poking when the fixed delivery-delay has passed without a pellet delivery. These plots are computer generated but they are best understood by imagining that there is a pen that traverses the paper horizontally from left to right on each trial, moving at an unvarying speed. When the mouse's head interrupts the infrared beam inside the hopper, the pen is "down," writing on the paper; when the head is not in the hopper, the pen is "up," not writing on the paper. Thus, the black lines show the intervals when the head was in the hopper; the white interruptions, the intervals when it was not. When food is delivered, a small circle is superposed on the black line. On food trials, the head is withdrawn soon after the pellet is released. On probe trials, there is no pellet and the mouse keeps its head in the hopper well beyond the expected time of pellet release. The variations in the well marked onset of poking may be automatically extracted to constitute the so-called start times statistic. The variations in the times of last head withdrawal on probe trials may be automatically extracted to constitute the so-called stop-times statistic. Together these statistics define the peak interval, the interval that brackets the time when the mouse expects food delivery. The+/+ mouse was a wild type C57/B6; the +/- and -/- mice were, respectively, heterozygous and homozygous CLOCK null mutants. These are in essence plots of the raw data; yet, they enable one to see at a glance that the differences between the strains are at best very small. They are all timing the pellet release with the same degree of accuracy and precision. (Cordes & Gallistel, 2008) Plots like these may be automatically generated while the testing is in progress.

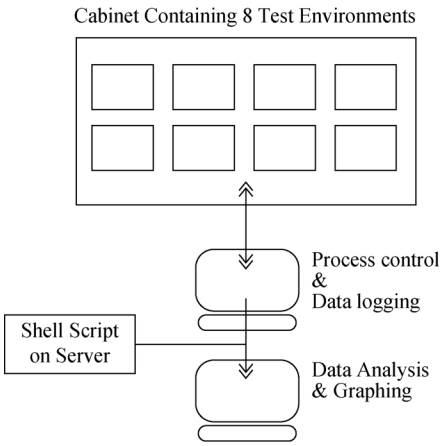

Figure 4. *Software schematic. Software written in the commercially available process-control language purchased with the mouse test chambers runs on a computer that controls and logs data from up to 8 test environments. A shell written in an object-oriented scripting language, running on a server, passes data periodically from the file into which the process-control software writes to the data-analysis software. The data analysis and graphing software analyzes the data to whatever level the user has specified in the creation of the data-analysis script. It communicates the results to the user by email. More extensive analyses may be performed off line at any time, using the same toolbox. The shell script archives the raw data when an experiment or a session is completed.*

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

At user-specified intervals, the shell copies the current version of the file to which the testing software is writing the data, and passes it to a data analysis script written with our Matlab toolbox. The data analysis script copies the data file into the Experiment structure, runs the scripted analyses, generates whatever graphs are specified within the analysis script, and, if so instructed, emails selected numerical results and selected graphs to a specified list of recipients. In this way, the progress of the testing is monitored at a high level of analysis without human intervention. We plan to provide for the reverse flow: The results of the almost real-time data analysis may be used to choose when and which test to run next, without human intervention.

This level of automation makes it possible in principle to screen hundreds of mice simultaneously. Early results from our fully-automated live-in environment suggest that our above list of test protocols aimed at our different target mechanisms may be completed in the span of 7-10 days. The principal cost is the one-time equipment cost, not ongoing salary costs. The equipment cost is not great in comparison to many other equipment costs that modern molecular biology labs must meet, nor in comparison to the cost of generating and maintaining mutant strains of mice.

The power and efficiency of the system is suggested by the following results from an illustrative sequence of three experimental protocols, run one immediately following the other, without any handling of the mice after their initial placement in the test environment. The sequence of 3 experimental protocols was run with on-line data analysis that enabled us to follow the progress of the experiment graphically in quasi real time. The results demonstrate that each mouse could: 1) Estimate its relative food incomes from two different food sources and adjust its relative temporal investments in these sources to match its relative incomes from them (1^{st} protocol, the "concurrent VI matching protocol"). 2) Learn to poke into one hopper in order to turn on a light inside one of two other hoppers—a short-latency hopper and a long-latency hopper—and learn to poke into whichever hopper lit up, in anticipation of food delivery (2^{nd} protocol, the "autoshaping" protocol, which assessed operant and Pavlovian conditioning simultaneously). 3) Learn to switch from a short-latency food source to a long-latency food source when the expected feeding latency at the short source passed without the delivery of a pellet (the "switch" protocol, which measures interval timing accuracy and precision). 4) Learn to anticipate the circadian time (time-of-day) at which the feeding hoppers would yield food (under one or the other of the just described "schedules or reinforcement"). The results from the first and third protocols (matching and switch) would also seem to imply intact spatial learning and memory: To respond appropriately, the mouse must remember which income or which feeding latency goes with which hopper; but, the two hoppers are identical; they are distinguished only by their position in the chamber, that is, by their spatial location. This rich yield of quantitative data on the functioning of basic mechanisms of cognition was obtained in a little more than one week, during which the mice were never handled, and during which a negligible investment of experimenter/technician time was required.

Six CB57BL/6 female mice were run simultaneously in six test environments contained in a single steel cabinet and controlled by a single computer running MedPC™ software. There was a 12:12 environmental light-dark cycle within the cabinet: light on 8:00-20:00; light off 20:00-8:00. The mice were placed in the test environments shortly before 16:00 of the first day. In order to test their ability to learn the times of day at which food was available, the to-be-described schedules of reinforcement were in force only from 21:00 to 23:00 and then again from 4:00 to 8:00 in each subsequent 24-hr period.

During the first 24-hr period, the protocol in force the two nightly foraging periods was a concurrent variable intervals (VI) protocol. A VI schedule delivers a food pellet in response to the first poke after a variable interval has elapsed following the previous release. In a concurrent VI

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

protocol, two VI schedules run independently on two feeding hoppers (see Figure 1). Mice move back and forth between the hoppers, poking first into one, then into the other, then back into the first, and so on. The relative incomes they obtain (pellets per unit time) are determined largely, but not entirely, by the expectations of the VI schedules. If the expected interval to the next pellet release at one hopper is twice that at the other, then the income from the first hopper will be roughly half of that from the second. Under these circumstances, animals (at least, vertebrates) adjust the average durations of their visits to the two food sources so that the ratio of the two expected visit durations approximately matches the ratio of the average incomes (pellets per unit time). This "matching" protocol measures the animal's capacity to estimate and remember the proportion (ratio) of two intensive magnitudes (amounts per unit time).

Real-time assessment of the accuracy with which a subject matches is obtained by plotting, on the same graph, pellet delivery by pellet delivery, the cumulative sums of the income and investment imbalances: *Income imbalance* = $(F_1-F_2)/(F_1+F_2)$, where F_i is the cumulative amount of food (number of pellets) obtained from Hopper *i*. It ranges from +1, when all the income has come from Hopper 1, to –1, when it has all come from Hopper 2. Similarly, *Investment imbalance* = $(T_1-T_2)/(T_1+T_2)$, where T_i is the cumulative duration of the visits to Hopper *i*. It, too, ranges from +1 to –1. The slope of a cumulative record is the average value of the successive measures being cumulated. Thus, during periods when the animal matches its investment imbalance to its income imbalance, the slopes of the two cumulative records are the same. Figure 5 shows these cumulative records for the 24 hours in which the matching protocol was in force.

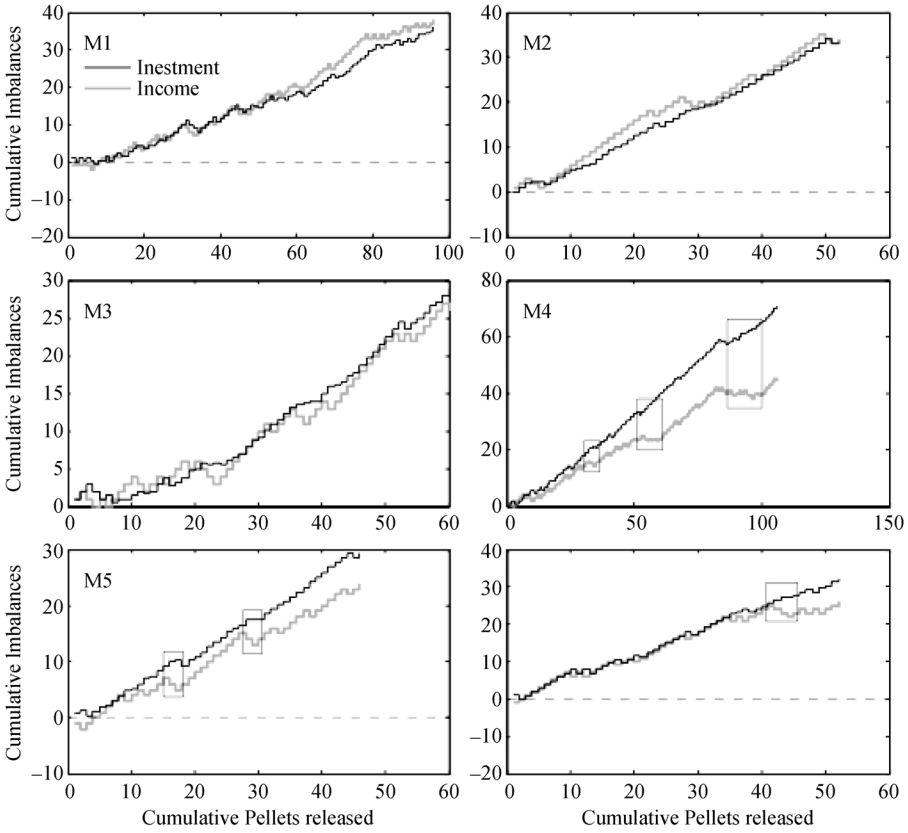

Figure 5. Pellet by pellet cumulative income and investment imbalances for 6 experimentally naïve mice during the first 24-hr period of a 3-protocol experiment. For the most part, the slopes of the two plots are the same. When they are, the mouse is matching the ratio of its average visit durations to the ratio of average incomes. Rectangles indicate periods when they are not.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

In accord with previously published results (C.R. Gallistel et al., 2007), one sees in Figure 5 that these ungentled, experimentally naïve mice matched from the outset. A single day-night cycle, with one 2-hour foraging period and one 4-hr foraging period, sufficed to determine that each mouse matched its investment ratio to its income ratio.

The test for matching behavior is a broad test of cognitive function: it requires the ability to estimate and remember temporal rates (number of pellets divided by elapsed time) and proportions (income ratios) and to program average visit durations so that the ratio of their average matches the ratio of the remembered incomes. Doing this requires the mouse to remember which hopper produces which income, and the hoppers are distinguishable only by their spatial location. It would seem that a mouse that could not remember could not match. Thus, this protocol provides an efficient screen for intact cognitive function, including memory function. With an 8-test-environment set-up, costing $40,000, one could screen 8*365 = almost 3,000 mice a year for intact cognitive and memory function, with a negligible demand on experimenter/technician time.

During the second 24-hour period, which followed immediately upon the first, the protocol in force assessed operant (instrumental) and classical (Pavlovian) conditioning. The foraging intervals, that is, the hours of the day when the mouse could obtain food from the two feeding hoppers, were the same as in the first protocol (21:00–23:00 and 4:00–8:00). This protocol introduced the illumination of a hopper as an information-bearing signal. The illumination of the trial-initiation hopper—located between the two feeding hoppers (see Figure 1)—signaled that the mouse could initiate a classical conditioning trial. The first poke into the illuminated trial-initiation hopper extinguished its light and illuminated one or the other of the flanking feeding hoppers. When the illuminated flanking hopper was the short-latency hopper (hereafter, the “short hopper”), a pellet was released into that hopper after 3s, regardless of the mouse’s behavior. When it was the long-latency hopper (hereafter, the “long hopper”), a pellet was released into that hopper after 9s, regardless of the mouse’s behavior. The long hopper was illuminated on a random 70% of the trials and the short hopper on the other 30%. The poke into the illuminated trial-initiation hopper is instrumental: it causes one or the other feeding hopper to light up and to deliver a pellet. The illumination of the trial-initiation hopper is a discriminative stimulus (SΔ), because it signals that a response (a poke into that hopper) will be effective. Therefore, learning to poke into the trial-initiation hopper is an instance of instrumental or operant conditioning.

By contrast, the release of a pellet into an illuminated feeding hopper is not contingent on the mouse’s behavior. Therefore, learning to poke into one of these hoppers in anticipation of pellet release is an instance of classical or Pavlovian conditioning. The illumination of a feeding hopper is a Pavlovian conditioned stimulus, because, as we will see, it elicits a conditioned response (poking into the illuminated hopper) only when it has repeatedly been paired with a reinforcement, such as the delivery of a food pellet. Instrumental and classical conditioning (aka operant and Pavlovian conditioning) are often thought to be mediated by different associative processes.

The cumulative record of the speed with which the mouse initiates a trial yields a measure of the progress of instrumental conditioning. Trial-initiation speed is the reciprocal of the latency between the illumination of the trial-initiation hopper and the first poke into after its illumination, which poke that initiates a trial. As may be seen in Figure 6, each mouse showed a more or less abrupt increase in trial-initiation speed at some point within the first 75 trials. The abruptness of the increase and the wide between-subject variation in when it occurs (ranging, in this case, from Trial 3 to Trial 72) accord with previously published findings (C.R. Gallistel, Balsam, & Fairhurst, 2004; R. W. Morris & Bouton, 2006; Papachristos & Gallistel, 2006).

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

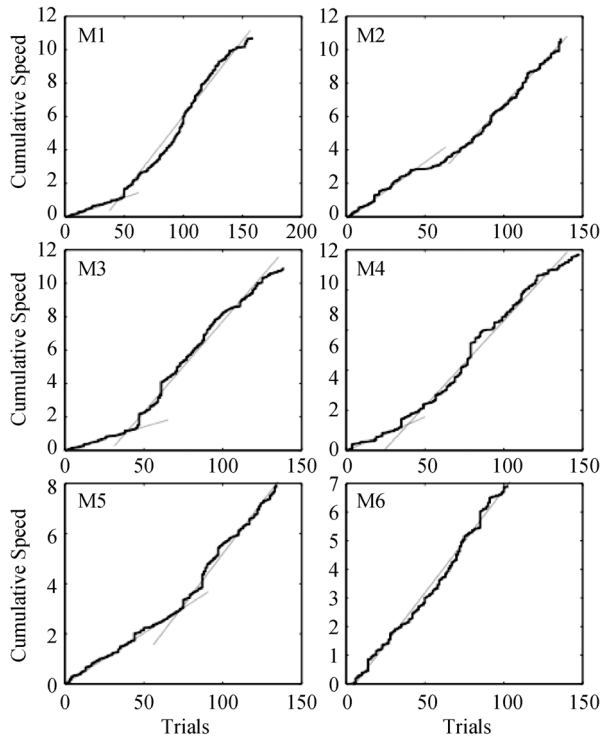

Figure 6. Trial-by-trial cumulative records of trial-initiation speeds (1/latency) for each of the 6 mice during the 2nd 24 hours of the 3-protocol experiment, when an instrumental and Pavlovian conditioning protocol was in force. Thin straight lines have been drawn to aid in the recognition of the change in slope that occurs when the speed with which trials are initiated increases more or less abruptly.

Similarly, the cumulative record of the difference between the poking rate into the illuminated feeding hopper during a trial (prior to the pellet delivery that terminates the trial) and the poking rate during the intertrial interval preceding a given trial yields a measure of the progress of the Pavlovian conditioning. As may be seen in Figure 7, the cumulative records of these differences in poking rates start out sloping downward. This means that the rate of poking during the illumination of the feeding hopper, when a pellet release is imminent, is less than the rate of poking into the unilluminated hopper during the intertrial interval, when a pellet release is not imminent. This is inhibitory effect of the unfamiliar CS is probably seen because the mice have learned during the preceding matching protocol to poke into the *unilluminated* feeding hoppers in search of food. Also, perhaps, because the illumination of these hoppers initially makes the mice wary and reduces their poking. However, in all but Mouse 2, there came a trial at which the slope of the cumulative record turned positive, indicating that the rate of poking during the CS (conditioned stimulus, hopper illumination) had become higher than the background rate. This trial, which may be objectively extracted from the data on CS-ITI rate differences using a change-point algorithm (C.R. Gallistel, Balsam, & Fairhurst, 2004), is the trial on which *that* mouse acquired the conditioned response to *that* CS. As was the case with the trials-to-acquisition measure for instrumental conditioning, and as is more generally the case with trials-to-acqusition (C.R. Gallistel, Balsam, & Fairhurst, 2004; R. W. Morris & Bouton, 2006; Papachristos & Gallistel, 2006), there was wide variation in this measure, across subjects and within subjects between CSs.

Trials-to-acquisition is a learning rate measure. The most commonly published measure of learning in the neurobiological literature is a trial-by-trial or block-by-block average of a behavioral measure, for example, average time to reach the platform in a water maze protocol or the average percent eye blinks in successive blocks of 10 trials in a conditioned eyeblink protocol. These group-average learning curves are generally taken to indicate learning rate, and between-group differences are uncritically taken to indicate between-strain differences in learning rate. Given the wide variations in trials-to-acquisition commonly seen when the data from each subject in a group are individually analyzed (C.R. Gallistel, Balsam, & Fairhurst, 2004; R. W. Morris & Bouton, 2006; Papachristos & Gallistel, 2006), given also the generally abrupt transitions seen in the individual curves, in marked contrast to the gradual changes in the group average plots (Papachristos & Gallistel, 2006), and given, finally, the strongly skewed distributions in trials-to-acquisition, one may doubt that group average learning curves reveal meaningful aspects of the underlying neurobiology. The group-average curves do not accurately describe the course of behavioral change in the

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

individual subjects. And, groups that differ in the mean may have strongly overlapping distributions, calling into question the reliability and physiological significance of this between-strain difference in the group-average learning curve.

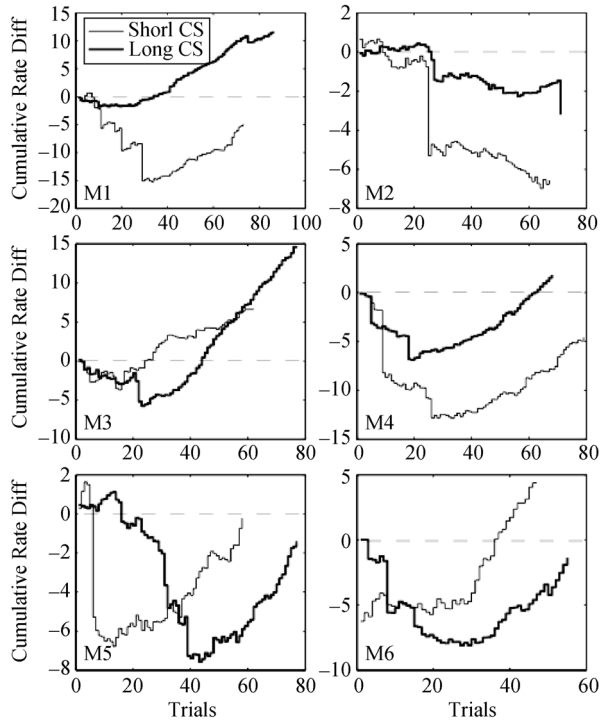

Figure 7. Trial-by-trial cumulative records of the difference in poking rate during the CS and during the preceding intertrial interval, from the 24-hour period when the instrumental and classical conditioning protocol was operative during the feeding phases. The trial at which the slope of the record becomes consistently positive defines the "trials-to-acquisition" for conditioned responding to a given CS. It marks the trial at which the mouse began to poke more frequently into that hopper in response to the CS (the illumination of that hopper) than during the intertrial interval (when the hopper was not illuminated).

The just-analyzed one-day phase of instrumental and classical conditioning laid the foundation for the third phase in which we used the switch protocol (Balci et al., 2008) to measure the accuracy and precision of interval timing. The instrumental conditioning trained the mice to initiate trials by poking into the illuminated middle hopper, while the classical conditioning taught them the feeding latencies associated with the illumination of each hopper (3 s for one hopper and 9 s for the other).

The third experimental protocol, the "switch" protocol, went into force when the mice had been in the test environment for 48 hours. In this protocol, as in the previous protocol, the mice initiated trials by poking into the illuminated trial-initiation hopper. In this protocol, the poke into the trial-initiating hopper illuminated *both* feeding hoppers, but only a randomly chosen one of them was set to deliver a pellet on that trial. The mouse had no way of knowing which hopper was set to deliver a pellet. On 30% of the trials, the short-latency hopper was set, while on the other 70%, it was the long-latency hopper. Moreover, 90% of all trials were operant trials rather than Pavlovian trials. On an operant trial, the pellet is released only in response to the first poke that extends through or occurs after the release-latency for the "set" hopper. By contrast, on a Pavlovian trial, the pellet is released at the release latency regardless of the subject's behavior. On operant trials, no pellet is released if the mouse pokes into the wrong hopper at the end of the release delay. If the long (9s) hopper is set and the first poke at or after 9s is into the short hopper, the trial ends without a pellet release. Likewise, if the short (3s) hopper is set and the first poke at or after 3s is into the long hopper, the trial also ends without a pellet release.

As with the previous two protocols, this feeding schedule was only operative during the hours 21:00-23:00 and 4:00-8:00. Unlike, the previous two protocols, which each ran only for one 24-hour period, this one ran for 5 such periods, partly to insure that switch performance has stabilized and partly because it was a weekend when the experiment ran unattended. At the end of the fifth 24-hr period, we began a 2nd switch session, which was just like the first switch session, except that now the probability of a short trial was .7 versus .3 for a long trial (reversing the relative probabilities that were in force during Session 3, the first of the switch sessions).

The mice began almost immediately to poke first to the short-latency feeding hopper and then, on trials where it did not deliver at the end of 3 s, to switch to the long latency hopper. The decision to leave the short-latency hopper depends on the

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

mouse's estimate of the time elapsed since the trial began, on memory for the duration of the two possible release latencies (3s and 9s) and on its estimate of the probability of a short- versus a long-latency trial (Balci, Freestone, & Gallistel, 2009). This latter estimate depends on its memory for the outcomes of a past sequence of trials.

The heavy curves in Figure 8 are the cumulative distributions of switch latencies from the last 100 trials in each session, for each of five mice in the two switch sessions. (The 6th mouse died at the beginning of the first feeding phase of Session 3, apparently from choking on a pellet). The thin vertical lines mark the two pellet-release latencies, 3s (on short trials) and 9s (on long trials). The probability that a mouse would leave the short hopper too soon or too late may be read directly from these cumulative distribution functions. The first probability of a premature departure is the value on the y-axis at which the cumulative distribution intersects the thin vertical line at 3s; the probability of leaving too late is 1 minus the value at which the curve intersects the thin vertical line at 9s. Consistent with previously published results (Balci, Freestone, & Gallistel, 2009), the great majority of the switches fall between the temporal goal posts at 3 s and 9 s. This implies that: i) they accurately remembers the two possible release latencies; ii) they remember which latency is associated with which hopper; iii) they can accurately compare the latency elapsed on any given trial to these remembered latencies; iv) based on that comparison, they can choose an appropriate target time for departing from the short-latency hopper. Moreover, and again in accord with previously reported results (Balci, Freestone, & Gallistel, 2009), the choice of a target departure time depends in an appropriate and approximately optimal manner on the relative probability of a short versus a long trial. When the relative probability shifts from .3:.7 in favor of a long trial to .7:.3 in favor of a short trial, the risk of suffering a pellet loss from a premature departure goes up, while the risk of suffering a pellet loss from a too-late departure goes down. Therefore, an optimal decision maker will shift the target departure time (hence the distribution of switch times) away from the short latency goal post and toward the long latency goal post. That is what every mouse in Figure 8 in fact did.

The distributions of switch times are approximately Gaussian, with two qualifications: 1) Our procedure censors the right tails of these distributions. Because a long trial ends without a pellet release when the mouse stays too long at the short hopper, the procedure prevents our observing switch latencies longer than 9s, which is what is meant by saying that the distributional data on departure latencies are right-censored. 2) As previously reported (Balci, Freestone, & Gallistel, 2009), some mice under some conditions impulsively switch to the long-latency hopper very early in the trial on some fraction of the trials. When these impulsive switches are frequent, the switch latencies form a bimodal mixture distributions (for clear examples, see the Session 3 plots for Mice 3 & 4 and the Session 4 plot for Mouse 1). We have found that the first component of these mixture distributions (the impulsive component) is well described by an exponential, while the second is well described by a Gaussian. In the light of these qualifications, we use Matlab's mle (maximum likelihood estimation) command to find the best fitting expgauss mixture distribution with the data censored at 9 s. The expgauss mixture distribution has 4 parameters: i) the time constant of the exponential component, ii) the mean of the Gaussian component, iii) the standard deviation of the Gaussian component, and iv) the relative proportions of the two components in the mixture. The best-fitting mixture distributions are the smooth thin curves superposed on the heavy empirical curves in Figure 8. They are only partially visible because they tend to fall directly on the empirical distributions. The likelihood-maximizing values for the means and standard deviations of the Gaussian component in each mixture are given at the bottom of each plot. The shift in the target departure time

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

produced by the change in the relative probability of a short versus a long trial is the difference between the two means. The smaller values for the sigmas in the lower row of plots show that this manipulation caused the mice to tighten up their switch latencies: in every case, the coefficient of variation, which is the proportion that the standard deviation bears to them mean, shrank between Session 3 and Session 4. In Session 4, this measure of timing precision was strikingly consistent from mouse to mouse:. 18, .17, .15, .17 and .18—in contrast to the highly variable trials-to-acquisition (learning rate) measure in Session 2.

As already noted, throughout the sequence of three protocols run over 9 consecutive 24-hour periods, beginning at 16:00 on the first day, the schedules of reinforcement (pellet delivery) only operated between 21:00–23:00 and 4:00–8:00. We restricted food availability to these recurring times of day so that we could get data on feeding-anticipatory poking activity/ Poking in anticipation of the onset of a feeding interval would suggest that our subjects also learned the circadian phases (times-of-day on their internal circadian clock) at which food became available in the test box.

In Figure 9, one sees that, after the first day, there is anticipatory poking in the 15 minutes prior to the onset of almost every feeding period. In an unpublished pilot experiment, we omitted one or the other feeding and observed heightened poking throughout most of the interval when food “should have” been available. Thus, while learning the contingencies in our schedules of reinforcement, the mice also learned to anticipate the onset of the feeding periods, either by timing the interval elapsed since lights out or by reference to their endogenous circadian clock. Given the literature (see Gallistel, 1990, for review), we believe that the latter is the more likely hypothesis. To prove that this food-anticipatory poking is based on a comparison between the current phase of a subject’s circadian clock and a remembered phase, we will need to provide food ad libitum and eliminate the light-dark cycle, so that the circadian behavioral

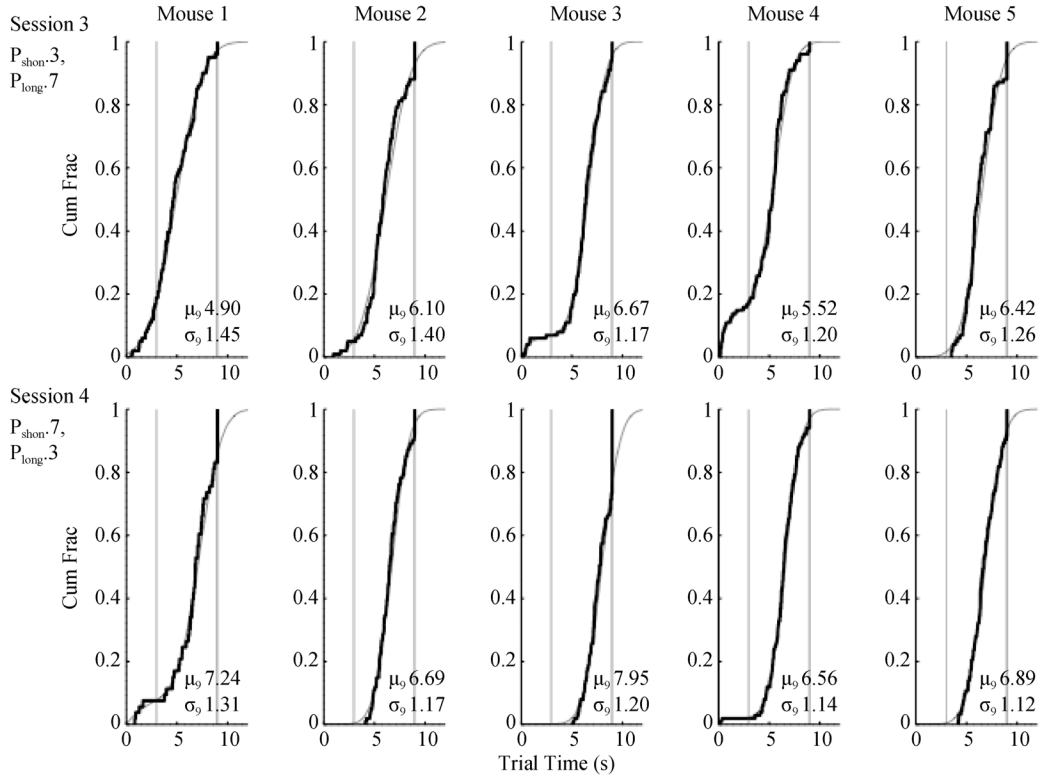

Figure 8. Cumulative distributions of switch latencies, with best-fitting expgauss mixture distributions superposed. The maximum likelihood estimates of the Gaussian mean and standard deviation are given at the bottom of each plot. The thin vertical lines mark the short and long pellet-release latencies. The relative probabilities of a short versus a long trial reversed between Session 3 and Session 4.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

activity clock runs free (unentrained by the light-dark Zeitgeber), then remove the ad libitum food and see whether anticipatory poking is seen at the appropriate phase of the subject's free-running clock.

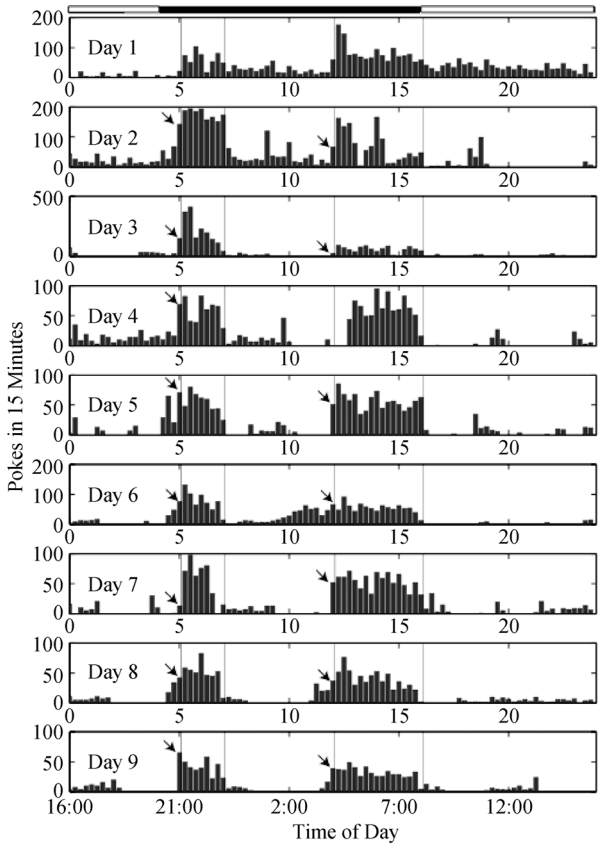

Figure 9. Representative bar plots of pokes in each quarter hour (15 minutes) throughout each day, for the 9 consecutive days of the experiment, for one mouse (Mouse 4). The thin vertical lines delimit the feeding periods, from 21:00-23:00 and 4:00-8:00. The light-dark cycle is shown at the top. Arrows point to anticipatory poking in the 15 minutes prior to most of the feeding periods after the first day.

Discussion

We stress the following aspects of our system for screening mutant, genetically manipulated and/or pharmacologically treated strains of mice:

- It gives physiologically meaningful quantitative parameters of well-characterized cognitive mechanisms. The quantities we extract could reasonably be compared to quantities that may now or someday be extracted from neurobiological methods operating at the cellular and/or molecular level of observation. It is the quantitative correspondence between behaviorally extracted and cellular and molecular measurements that carry the greatest conviction when it comes to linking behaviorally defined mechanisms to the underlying neural and molecular biology.
- It measures the *contents* of memory rather than simply the existence of a memory of some kind. Our methods measure the accuracy and precision with which mice remember relative incomes, intervals, probabilities, proportions, and times of day. These measures of accuracy and precision are to a great extent independent of variations in levels of overall performance. The numbers of trials initiated varied considerably between our mice, as did the overall amount of poking during a trial, but the parameters of their matching and timing were very similar. In a given feeding period, one mouse may make very few switches, while another makes many switches, but the timing of the switches they do make is highly comparable in its accuracy and its precision.
- It is highly efficient in its consumption of temporal, spatial and human resources. The rich results that we have just summarized were obtained over a span of 9 days, with mice that were ungentled and experimentally naïve when placed in the test environment; they could probably have been obtained in 2 or 3 fewer days than that. A cabinet containing eight 22"×20" ×11" (53×48×26 cm) environments has a 48" ×24" (1.15 × .53 m) footprint. Several cabinets may be placed in an ordinarily sized experimental room. Our results were obtained with an almost negligible investment of human time. If we do not count the time spent studying the plots, the only time required to produce the different results was the time to place the mice in the test environments at the beginning of the 9 days and the 5 minutes at the computer console required to change from one protocol to the next.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

- An enormous amount of time has gone into the development of the software, both the custom Matlab toolbox, which makes it easy for those with modest computer programming skills to do complex analyses, and the code that uses those commands to do the analyses and plots here reported and shown. However, that code is now publicly available, along with extensive tutorial material on its use. Thus, the highly-informative suite of 3 protocols here described can be run as a turn-key operation, using publicly accessible software for both control and data analysis.
- The software produces results and graphs of results in quasi real time. The experimenters know what the mice are doing within hours of their doing it. One can study the graphs of the results from the previous night's feeding phases over one's breakfast coffee.
- Our system eliminates the handling of the mice, greatly reducing the stress they experience during the behavioral testing. This is good for the mice, and good for the data.
- The test environment is an enriched environment, considerably more like a natural environment than is a simple tub. The mice are not in any serious sense food deprived. They get 6 hours access to food in each 24 hours, one an hour before and after dusk and one four hours before dawn. This roughly simulates nocturnal rodent feeding patterns in the wild. There is no need to weigh them, which is itself a stressful procedure. We know their food intake, because the system monitors the arrival and removal of each released pellet, by means of an IR beam at the bottom of the V-shaped trough in each feeding hopper.
- An unexpected benefit is that the mice urinate and defecate mostly in the test box, which has a removable feces tray beneath a grid floor. This natural inclination to relieve themselves outside the nest leaves the bedding in the nest tub clean. Veterinary staff should encourage this form of behavioral testing as being more humane than the warehousing of mice in tubs, the daily weighing, and the allocation of a daily ration.

References

Antle, M. C., & Silver, R. (2005). Orchestrating time: arrangements of the brain circadian clock. *Trends Neurosci, 28*(3), 145–151.

Balci, F., Allen, B. D., Frank, K., Gibson, J., Gallistel, C. R., & Brunner, D. (2009). Acquisition of timed responses in the peak procedure. *Behavioral Processes, 80*, 67–75.

Balci, F., Freestone, D., & Gallistel, C. R. (2009). Risk assessment in man and mouse. *Proc Natl Acad Sci U S A, 106*(7), 2459–2463.

Balci, F., Papachristos, E. B., Gallistel, C. R., Brunner, D., Gibson, J., & Shumyatsky, G. P. (2008). Interval timing in the genetically modified mouse: A simple paradigm. *Genes, Brains & Behavior, 7*, 373–384.

Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends Neurosci, 32*(2), 73–78.

Bruce, V. G., & Pittendrigh, C. S. (1956). Temperature Independence in a Unicellular "Clock". *Proc Natl Acad Sci U S A, 42*(9), 676–682.

Church, R. M. (1984). Properties of the internal clock. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Vol. 423, pp. 567–582). New York: New York Academy of Sciences.

Cordes, S., & Gallistel, C. R. (2008). Intact interval timing in circadian CLOCK mutants. *Brain Res, 1227*, 120–127.

Dubnau, J., Chiang, A. S., & Tully, T. (2003). Neural substrates of memory: from synapse to system. *J Neurobiol, 54*(1), 238–253.

Eichenbaum, H. (1996). Learning from LTP: a comment on recent attempts to identify cellular and molecular mechanisms of memory. *Learn Mem, 3*(2-3), 61–73.

Foster, R. G. (1993). Photoreceptors and circadian systems. *2*, 34–39.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.

Gallistel, C. R. (1999a). Coordinate transformations in the genesis of directed action. In *Bly, Benjamin Martin (Ed); Rumelhart, David E (Ed) (1999) Cognitive science Handbook of perception and cognition* (2nd ed) (pp 1–42) xvii, 391pp.

Gallistel, C. R. (1999b). The replacement of general-purpose learning models with adaptively specialized learning modules. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences. 2nd ed.* (2 ed., pp. 1179–1191). Cambridge, MA: MIT Press.

Gallistel, C. R. (2002). Frequency, contingency and the information processing theory of conditioning. In *Sedlmeier, Peter (Ed) Dept of Psychology; Chemnitz U of Technology; et al (2002) ETC Frequency processing and cognition* (pp. 153–171). London: Oxford University Press.

Gallistel, C. R. (2003). Conditioning from an information processing perspective. *Behav Processes, 62*(1-3), 89–101.

Gallistel, C. R. (2007). Flawed foundations of associationism? Comments on Machado and Silva (2007). *Am Psychol, 62*(7), 682–685; discussion 689-691.

Gallistel, C. R. (2008). Learning and representation. In R. Menzel (Ed.), *Learning theory and behavior. Vol 1 of Learning and memory: a comprehensive reference* (Vol. 1. Learning theory and behavior, pp. 227–242). Oxford: Elsevier.

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proc Natl Acad Sci U S A, 101*(36), 13124–13131.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychol Rev, 107*(2), 289–344.

Gallistel, C. R., & King, A. P. (2009). *Memory and the computational brain: Why cognitive science will transform neuroscience*. New York: Wiley/Blackwell.

Gallistel, C. R., King, A. P., Gottlieb, D., Balci, F., Papachristos, E. B., Szalecki, M., et al. (2007). Is matching innate? *J Exp Anal Behav, 87*(2), 161–199.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Gallistel, C. R., Shizgal, P., & Yeomans, J. S. (1981). A portrait of the substrate for self-stimulation. *Psychol Rev, 88*(3), 228–273.
Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Vol. 423, pp. 52–77). New York: New York Academy of Sciences.
Kandel, E. R. (2004). The molecular biology of memory storage: a dialog between genes and synapses. *Biosci Rep, 24*(4-5), 475–522.
Koch, C. (1999). Biophysics of Computation: Information Processing in Single Neurons. *Oxford University Press, Oxford*.
Leil, T. A., Ossadtchi, A., Cortes, J. S., Leahy, R. M., & Smith, D. J. (2002). Finding new candidate genes for learning and memory. *J Neurosci Res, 68*(2), 127–137.
Markram, H., Lubke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science, 275*(5297), 213–215.
Maywood, E. S., O'Neill, J. S., Reddy, A. B., Chesham, J. E., Prosser, H. M., Kyriacou, C. P., et al. (2007). Genetic and molecular analysis of the central and peripheral circadian clockwork of mice. *Cold Spring Harb Symp Quant Biol, 72*, 85–94.
Mercer, T. R., Dinger, M. E., Mariani, J., Kosik, K. S., Mehler, M. F., & Mattick, J. S. (2008). Noncoding RNAs in Long-Term Memory Formation. *Neuroscientist, 14*(5), 434–445.
Morris, R. G., Moser, E. I., Riedel, G., Martin, S. J., Sandin, J., Day, M., et al. (2003). Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory. *Philos Trans R Soc Lond B Biol Sci, 358*(1432), 773–786.
Morris, R. W., & Bouton, M. E. (2006). Effect of unconditioned stimulus magnitude on the emergence of conditioned responding. *J Exp Psychol Anim Behav Process, 32*(4), 371–385.
Papachristos, E. B., & Gallistel, C. R. (2006). Autoshaped head poking in the mouse: a quantitative analysis of the learning curve. *J Exp Anal Behav, 85*(3), 293–308.
Pittendrigh, C. S. (1960). Circadian rhythms and the circadian organization of living systems. *Cold Spring Harb Symp Quant Biol, 25*, 159–184.
Pittendrigh, C. S. (1965). On the mechanism on the entrainment of a circadian rhythm by light cycles`. In J. Aschoff (Ed.), *Circadian Clocks* (pp. 277–300). Amsterdam: North-Holland.
Quinn, W. G. (2005). Nematodes learn: now what? *Nat Neurosci, 8*(12), 1639–1640.
Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 1–46). New York: Academic.
Richter, C. P. (1965). *Biological clocks in medicine and psychiatry*. Springfield, Ill: Thomas.
Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal, 27*, 379–423, 623–656.
Shors, T. J., & Matzel, L. D. (1997). Long-term potentiation: what's learning got to do with it? *Behav Brain Sci, 20*(4), 597–614; discussion 614–555.
Takahashi, J. S. (2004). Finding new clock components: past and future. *J Biol Rhythms, 19*(5), 339–347.
Thompson, S. M., & Mattison, H. A. (2009). Neuroscience: Secret of synapse specificity. *Nature, 458*(7236), 296–297.
Weiner, J. (1999). *Time, love, memory: A great biologist and his quest for the origins of behavior*. New York: Alfred A. Knopf.
Wood, M. A., Kaplan, M. P., Park, A., Blanchard, E. J., Oliveira, A. M., Lombardi, T. L., et al. (2005). Transgenic mice expressing a truncated form of CREB-binding protein (CBP) exhibit deficits in hippocampal synaptic plasticity and memory storage. *Learn Mem, 12*(2), 111–119.
Zhang, C. L., Zou, Y., He, W., Gage, F. H., & Evans, R. M. (2008). A role for adult TLX-positive neural stem cells in learning and behaviour. *Nature, 451*(7181), 1004–1007.

一种对动物学习记忆能力突变筛查的新途径

C. R. Gallistel[1], A. P. King[2], A. M. Daniel[1], D. Freestone[1], E. B. Papachristos[1], F. Balci[1], A. Kheifets[1], 张靖 [1,3], X. Su[1], G. Schiff[1], H. Kourtev[1]

([1] 罗格斯大学心理学系认知科学研究中心, 新布伦士维克, 新泽西州, 美国)

([2] 费尔菲尔德大学数学计算机科学系, 费尔菲尔德, 康乃狄格州, 美国)

([3] 中国科学院心理研究所行为遗传学研究中心, 北京 100101)

摘 要 本文详细描述了一种全自动化的行为检测方案。在自然饲养环境/检测环境(24/7)中, 我们测量了小鼠针对两个给食器中获得食物的比例与它们在相应给食器停留时长的比例进行匹配的精确性与准确性。该方案是对传统条件性习得 (trials-to-acquisition) 行为测验设备的改进, 可以检测动物时间间隔能力的精确性与准确性, 对定时目标选择的相关概率的效果, 以及记忆一天中从不同给食器中获得食物次数的精确性与准确性。该压缩系统避免了在整个实验过程中对小鼠的持握操作, 可忽略实验者/技术员的实验操作时间, 而且可以递送小鼠置入实验环境后, 7~9 个实验日中全部 3 组实验流程产生的大量结果。其中, 第一个实验流程为单个 24 小时周期内完成的时间匹配能力的筛查, 它对动物的时间、空间估计能力的记忆机制进行精确检测。因此, 该系统允许在有限的实验空间、较短的实验周期内, 对大量的实验小鼠进行有可能存在的学习记忆能力缺陷进行大规模筛查。此外, 该系统运行所依赖的软件可以在公共开放平台获得。

关键词 时间记忆; 匹配; 表型筛选; 突变; 自动化

分类号 B841; B845

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net