SEE: a tool for the visualization and analysis of rodent exploratory behavior

Dan Drai, Ilan Golani*

Department of Zoology, George S. Wise Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv, Israel

Abstract

The complexity of exploratory behavior creates a need for a visualization and analysis tool that will highlight regularities and help generating new hypotheses about the structure of this behavior. The hypotheses can then be formulated as algorithms that capture the patterns and quantify them. SEE is a Mathematica based software developed by us for the exploration of exploratory behavior. The raw data for SEE are a time series of the animal 's coordinates in space sampled at a rate that allows a meaningful computation of speeds. SEE permits: (i) a visualization of the path of the animal and a computation of the dynamics of activity; (ii) a decomposition of the path into several modes of motion (1st gear, 2nd gear, etc.) and a computation of the typical maximal speeds, the spatial spread, and the proportion of each of these modes; and (iii) a visualization of the location in the environment of stopping episodes, along with their dwell time. These visualizations highlight the presence of preferred places, including the animal's so-called home base, and permits a computation of the spatio-temporal diversity in the location of stopping episodes. The software also: (i) decomposes the animal's path into round trips from the home base, called 'excursions', and computes the number of stops per excursion; (ii) generates a visualization of the phase space (path + speed, traced in a three-dimensional graph) of any progression segment or list of such segments; and (iii) produces a visualization of the way places in the animal's operational world are connected to each other. SEE also permits the definition and computation of behavioral endpoints across any section of any database of raw data. The range of applicability of SEE to various experimental setups, tracking procedures, species, and preparations is addressed in the discussion. © 2001 Published by Elsevier Science Ltd.

Keywords: Phenotyping mouse behavior; Locomotor behavior; Computerized tracking; Spatial memory; Open field; Measuring behavior; Cognitive map; Behavior genetics; High throughput phenotyping; Place fields

1. Introduction

The view that rat open field behavior is largely stochastic (e.g. [26]) gave way recently to the realization that it is quite patterned. A series of studies have shown that this behavior includes several ethologically-relevant precisely definable and measurable patterns: when placed in a novel environment a rat establishes a home base from which it performs round-trips (excursions) into the environment [20]. These excursions gradually grow in amplitude, having a slow and intermittent outbound portion and a fast and continuous inbound portion. With extended exposure, this velocity profile changes in a predictable manner [66]. The rat's trajectory in the environment consists of several modes of motion (1st, 2nd and 3rd gears); progression without leaving 1st gear (lingering) [16] defines what has been previously labeled 'stopping behavior', and the number of stops per excursion is constrained by an upper bound [32].

Aspects of this structure have been uncovered in long-term studies of tame wild rats, hooded laboratory rats in a variety of environments and contexts, and drug-treated preparations [21,22,23]; it has also been recently applied to the phenotyping of mouse behavior [5,17]. To establish a comparative study of this complex structure, it was necessary to develop a special tool of visualization and analysis. This tool would highlight regularities, help generating new hypotheses, and allow statistical inferences about the structure of the behavior. While describing this tool, this paper also provides the reader with a working environment for exploring his own data of rodent behavior. A Mathematica notebook version of this paper, including visualizations and animations of mouse behavior, can be downloaded at http://www.tau.ac.il/~ilan99/

Recent studies demonstrating the vulnerability of behavioral measures to environmental manipulations [12] have aroused an interest in ethologically relevant

parameters of rodent behavior. This is because such parameters might be more resistant to environmental manipulations. The interest in ethological parameters in behavior genetics [29], the need for animal models with high ecological validity in ethopharmacology [60], and the need for high throughput phenotyping of mouse behavior [52] all coincide with our own interest in developing an algorithmic approach to the ethological study of rodent behavior. The present study is one in a series introducing this approach. As such, it is motivated by a search for intrinsic constraints, which should hopefully show relative stability across experimental protocols and, thus, provide more reliable measures in the study of the genome/behavior interface. Being ethologically motivated, this study does not assume a priori constructs such as fear, anxiety, learning, or even investigation (although our software can readily be used within paradigms involving these constructs). Rather, it lets these and other constructs emerge as 'by-products' of lower level movement processes. We suspend judgment about meaning *because* our main motivation has been to reveal the function and meaning of behavior on the basis of its structure [33]. The traveling speed of a rat embodies its orientation with regard to its home base: slow progression indicates progression away from home, and fast progression indicates going back home [65]. An upper bound on the number of stops per round-trip from home embodies some form of spatial memory [32], etc. Other features visualized in this study similarly embody a variety of cognitive and motivational properties of the behaving animal.

In the first paper in this series [16], we established the algorithmic definition and the subsequent classification of rodent locomotor behavior into three modes of motion. In the present paper, we use this classification to present some of the visualization tools developed by us, and describe their software implementation, which we call SEE (a Software for the exploration of exploration). We also show how this low-level classification generates hypotheses about how the animal occupies the environment, how it connects its various parts, how it distinguishes between inbound and outbound trips, etc. The third paper in this series is dedicated to the use of statistical inference for the establishment of behavioral endpoints that differentiate between two mouse strains [18,19].

The level at which the behavior is analyzed is that of the whole animal considered as a moving point. The primary data are a time series of the rat's coordinates in space, sampled at a rate that permits a meaningful computation of speeds. These data, obtained by an automatic tracking system, are the input of our software environment.

The first section of this paper is devoted to a review of some of the already established patterns of rat exploratory behavior. These patterns provide the articulated framework within which SEE is implemented. The second section illustrates how these patterns and others can be visualized, then formalized, and finally measured with the help of SEE. The

statistical tools for comparing these measurements across strains, species, or preparations are described elsewhere [5,17,18].

In the third section, we discuss the potential use of this approach in behavioral neuroscience and behavior genetics. An Appendix provides a formal description of the software (specification of data objects and operators).

The full implementation of the software is available on request from the authors.

2. Methods

2.1. The programming environment

The Mathematica software [74] mediates the interaction with SEE. Mathematica is both a programming language and a graphical display tool. One can write and save function definitions, and use them afterwards in interactive Mathematica sessions. SEE is, thus, nothing but a collection of function definitions, written in the Mathematica programming language. A prerequisite to the use of SEE is the smoothing of the data, the computation of speeds and the computation of the segmentation of the behavior into modes of motion [16].

Using SEE consists, therefore, of loading the appropriate function definitions and data, and then interactively visualizing the data or computing numerical attributes of the data by invoking a pre-defined function. In a more advanced mode, the user can also add new function definitions as the need arises. This is the main advantage of working in such an environment: it enables an easy and unlimited extension of the capacities of SEE, as new ideas about the structure of the data (exploratory behavior) emerge.

After presenting the main patterns that were revealed so far, we will focus on one session of an animal (rat #3, fifth session) from an experiment with hooded rats in which the animals were exposed repeatedly to a large arena (6.5 m diameter) for sessions of half an hour [66]. A few computations or graphics pertain to all the eight daily sessions of that animal. In order to give the reader an opportunity to practice the syntax and the (full) commands that generate each of the graph types and/or calculations, animations, etc., we present a working version of this study on the web (see the Introduction).

3. Background on the structure of exploratory behavior

3.1. Stops

Anyone can see that when a rat or a mouse is placed in a novel environment it alternates between progression and stopping. Each time the animal stops it performs scanning movements—sniffing, establishing snout contact with the substrate, and/or looking around. The type, rate, and number of scans presumably determine the amount and type of

information gathered by the animal. Forward progression carries the animal from one location to the next, while stopping and scanning involve investigation of particular locations [6]. These two patterns constitute locomotor behavior, a *par excellence* innate pattern [44]. Only recently has it been shown that there is an inverse relationship between these patterns also at the level of the CNS, and that this relationship is mediated by the hypothalamus [62]. Clearly, these patterns serve separate and distinct functions, and it is only reasonable to expect that measuring and quantifying them separately would help differentiate between species, strains, and preparations.

Because in rodents stopping largely implies scanning, and scanning implies cessation of progression and stopping, measuring stopping entails an indirect measurement of scanning, which in turn implies measurement of information acquisition and attention-involving processes. The type of a scan determines the duration and the spatial spread of a stopping episode: a forward scan involving only the head and neck is typically shorter in duration than that of a scan recruiting the whole body, as in 'stretched attention'. While performing these scans, the animal's center of gravity changes its location; the spatial spread of this change, i.e. the length of the path traced by the animal on the ground while performing the scan, is small for head scans and large for whole body scans that may even involve a step or two. The number of scans performed during a stopping episode also affects the duration and the spatial spread of a stopping episode. All these ethologically meaningful patterns influence the duration and spatial spread of stopping episodes. Capturing and quantifying their spatio-temporal features should, therefore, prove a fruitful phenotyping tool. When behavior geneticists score rearings (e.g. [13]), and ethopharmacologists score stretched attention episodes (e.g. [60]), they capture only the full-blown scans along the respective vertical and forward so-called collective variables [33], and lose the vast majority of smaller amplitude scans along the vertical, forward, and horizontal dimensions. Practically all of these scans are captured, if only in an indirect and pooled way, in SEE.

Until recently stopping was defined as a classical behavior pattern based on the observer's subjective judgment. In some studies, the definition of a stop was based on an arrest of the hind legs [32], a criterion that is difficult to measure automatically. In other studies, it was based on an arbitrary (though reasonable) threshold value for velocity [11,30], e.g. if velocity is lower than 2 cm/s call it a stop. In still other studies 'the circumscribed movements involving the processing of localized sensory stimuli' were captured by characterizing the overall geometry of the path [55].

Capturing individual stops on the basis of intrinsic 'natural' criteria is not a straightforward task. This is because (i) stopping is not to be equated with arrest, i.e. it does not necessarily consist of zero speed and (ii) the cutoff speed values distinguishing between progression and 'stopping' differ from species to species and from strain to strain.

To distinguish between stopping and progression, we have to estimate speeds, i.e. the first derivative of the (noisy) location data. The derivation further increases the level of the noise so that smoothing the data becomes even more essential. On the other hand, because of their ethological significance, we do not want to lose stops (which are sometimes as short as 0.2 s) in the process of smoothing the data. For a video clip illustrating the 'smeared' nature of stopping behavior see http\\www.tau.ac.il~ilan99

The steps for establishing the cutoff point between progression and 'stopping' (or lingering episodes, as we call them), are rather involved, but essential (a user-friendly stand-alone software package establishing the cutoff point is available from the authors). The software performs the following steps: (i) it establishes the noise level of the system; (ii) it defines 'sub-noise' periods as periods of 'arrest'; (iii) it uses arrest periods to segment the velocity time series into 'motion segments' ('inter-arrest' intervals); (iv) it establishes the speed maxima of each of the motion segments (we previously used a corresponding measure of the local spread of motion called 'local movement'— 'locmov' [16]; (v) for each rat-session it computes the frequency distribution of the speed maxima (or 'locmov') of all its motion segments; (vi) it uses a Gaussian mixture model to analyze the distribution of these maxima. This model is commonly used in electrophoresis, for example, for recognizing distinct components within a mixture (see Section 1.3.3. in the Appendix); (vii) it estimates the parameters of the Gaussian model by using the Expectation-Maximization (EM) algorithm [15]; For a nice exposition see [25]. This algorithm estimates the maximum likelihood parameters (proportions, means, and standard deviations) of a mixture with a given number of Gaussians (Section 1.3.3 in the Appendix). Having the Gaussians at hand, it can finally establish the thresholds that distinguish between them. Most important, it establishes the threshold that distinguishes between the leftmost Gaussian, representing the progression mode with the lowest maximal velocities. It is this component which was termed 'stopping' in our previous studies and is now termed 'lingering', 'staying in place behavior', or '1st gear mode'; and, finally (viii) bouts in which the animal alternates between full arrests and lingering episodes are joined into a single episode of lingering behavior. Therefore, at one extreme, a lingering episode may consist of an extremely short arrest. At the other extreme, it may consist of a long bout of low speed segments that alternate with long and/or short arrests. From our point of view, both extremes represent single episodes of staying in place behavior. Strains and species differ greatly in the consistency of lingering episodes. In the hooded rat whose behavior is illustrated in this paper, the threshold speed distinguishing between lingering and progression segments was $16 \pm 1.99(SD)$ cm/s (based on eight successive daily 1/2 h sessions in a 6.5 m diameter arena. In comparison, the Balb/cJtau mice median of maximal speeds during lingering episodes, based on a single session of eight animals in a

Cumulative time (of lingering episodes) per location
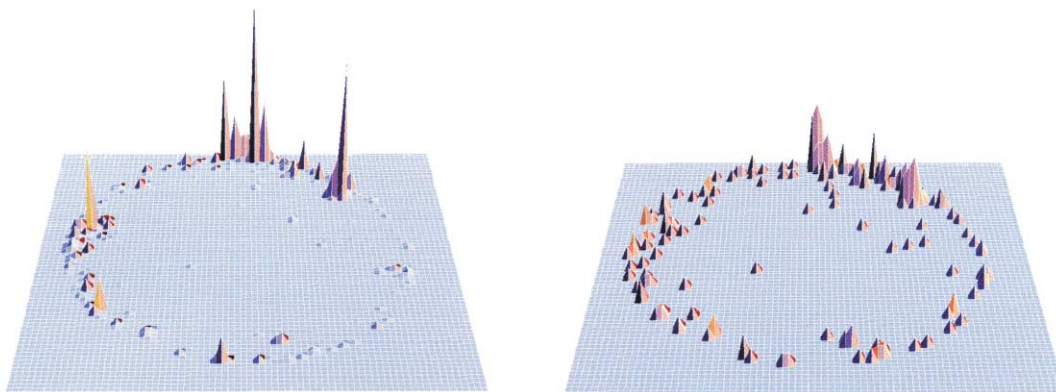
Number of visits (lingering episodes) per location

Fig. 1. The height of the graph at a particular location in the circular 6.5 m diameter arena represents: left—cumulative dwell time during lingering episodes (stopping) at that location. Right: number of visits (lingering episodes, stops) to that location. As shown, one neighborhood in the arena, at approximately 12 o'clock stands out from all the other places in the arena in the magnitude of these two measures. This neighborhood is defined as the animal's home base (here, there is also a 'garden', typically appearing when there are multiple sessions, at 1.00 o'clock).

3.2 m diameter arena, was found to be 3.35 cm/s (2.61, 4.30) cm/s, whereas that of a corresponding group of C57BL/6Jtau mice was 5.55 cm/s (4.53, 6.74) cm/s.

3.2. Movement segments

From the categorization of modes of motion described above, one can obtain a segmentation of the behavior into intervals of lingering and intervals of progression or movement segments. The path traced during a progression episode may embody a variety of sensorimotor and perhaps also cognitive algorithms used by the animal. Golani et al. [32] showed that hooded rats appear to scale inter-stop distances according to the size of the arena: when they encircle the arena with a single excursion they partition that excursion into a more or less similar number of segments, be it a 2×2 m or an 8×8 m arena. This means that the rat has some global view (see Ref. [28]) of the whole arena which determines the size of individual segments; it is not a particular distance that determines when a progression segment will end, but rather a particular fraction of arena circumference.

The long non-stop segments performed by hooded rats during inbound progression segments to home base (Fig. 6, 3rd from left) reveal the attraction exerted by the home base on the animal. A smooth ballistic speed trajectory of a movement segment (see in Fig. 10b, trajectory across the open space) betrays the fact that the segment has been pre-programmed by the animal at the outset of the movement (in contrast to an intermittent one (foreground trajectory in Fig. 10b), in which the animal, as it were, feels its way as it proceeds along the arena wall).

Peak velocities of movement segments range in the presently analyzed rat-sessions between a mean threshold of 16 cm/s ($M - SD = 14$, $M + SD = 18$) cm/s and a mean maximum of 286 cm/s (236, 336) cm/s.

3.3. The home base

When confronted with a novel environment, rats establish a preferred place, which they appear to use as both a cognitive and a motivational anchor. Repeated visits to the home base are used for grooming, a pattern considered to reduce arousal [14]. These repeated visits perhaps help the rat manage its level of arousal in a novel and presumably stressful situation. By measuring the rat's movements in relation to the home base, one comes to realize that the seemingly haphazard roaming of the rat is in fact systematic and organized.

The rat's preference for one (or in some cases two) place(s) can be shown by considering the spatial distribution of stops in the arena. In the preferred place, the cumulative duration of staying in place (whether we consider cumulative dwell time of stops or total cumulative time including both progression and stopping) is of a higher order of magnitude compared to the values scored in all the other places (Fig. 1, left). The number of stops is also the highest (Fig. 1, right), and the incidence of grooming episodes is significantly higher than expected by the higher proportion of time spent in that place. This place is termed the rat's home base [20].

In an arena devoid of any objects or places that suggest shelter, each rat establishes its own home base in a different place, early in the session. This suggests that home base location involves the use of spatial memory, and is not selected de novo each time the rat returns to it. If, however, one object that is clearly more familiar is present, most rats establish their home base in that place. This property has been used to standardize home base location, thus simplifying the comparison across animals [63,65]. Interestingly, home base location relative to the edge of the arena has been used by Cools et al. [10] as one of the features distinguishing between Apomorphine-Susceptible and

Apomorphine-Unsusceptible selected sub-populations of hooded rats.

3.4. The excursion

Using the location of the home base as a reference point, it is readily seen that Exploratory Behavior, at least in its initial stages, can be viewed as consisting of roundtrips (excursions) performed from the home base. Home base behavior (including excursions) was previously reported in rats [1,2,8,20]. It has also been reported in chimpanzees constructing a cognitive map of a novel arena in a zoo [46], in infant wild chimpanzees who study the environment and gain their independence while performing increasingly longer excursions away from their roaming mothers [69], and in bulls having a preferred place in the arena, where they also go to die after having been fatally injured in the course of a bullfight [47].

When hooded rats were exposed daily to the same large (6.5 m diameter) arena, excursion length increased both within and across daily sessions. While each animal had its own initial excursion length, all animals shared the same rate of excursion growth, both within and across sessions. This growth pattern could reflect a strain-specific capacity to handle the novelty of the environment [65]. A similar pattern of growth, involving the performance of increasing loops around a point of origin has been reported in desert isopodes (*Hemilepistus reaumuri*) [36–38], in spiders [49], and in desert ants (*Cataglyphis* sp.) [51,70]. Its presence in both invertebrates and vertebrates might reflect a fundamental adaptive cybernetic principle characterizing innate goal directed activities. While it could embody a familiarization process (a hypothesis examined systematically in an empirical [65] and a simulation study [66]) it could also develop on an idiothetic basis by repeating the same scheme over and over, each time extending its reach further away from the starting place. This possibility is supported by studies made by McNaughton et al. [45].

In the rat studies, the excursion was defined as a roundtrip that starts at the home base and ends, usually after the performance of one to several stops, also at the home base. Only the number of stops per excursion was scored (ignoring excursion mileage or duration). Using two robust statistical methods—density estimators and quantile plots—it has been shown that the number of stops per excursion fits a model of uniform distribution with an upper bound, rather than a model of normal or a geometric distribution [32]. This means that the probability of returning to the home base increases after each additional stop until it equals 1 at the upper bound. This supports the hypothesis that home base location depends on memory, and returning to it is not elicited de novo at the end of each excursion. Moreover, as mentioned briefly earlier, by increasing the size of the arena from 2×2 to 8×8 m, this upper bound of stops could not be increased. Instead, the rats cover the larger arena by scaling up the inter-stop

distances. The distribution of the rate of stopping per time and/or distance unit may account for many of the measured properties of Exploratory Behavior. For example, Whishaw et al. [71] have shown that the hyperactivity attributed to rats with Fimbria–Fornix lesions is not due to travel distance and speed, but rather due to the presence of significantly more stops of shorter duration (and, thus, more movement episodes). The distribution of the number of stops per excursion can be modified by the psycho-stimulant drug D-amphetamine [20], and has also been used to differentiate between the response of two strains of rats to amphetamine [11,30].

3.5. Speed profiles

The momentary traveling speed of an animal appears to both reflect and influence the type of perceptual input managed by it at that moment. This can provide one with an insight into its operational world [68]. When, for example, water shrews (*Sorex palustris*) kept in a terrarium on Lorenz's desk were alarmed, they rushed back home at top speed along the digressive much longer 'well trodden path', instead of taking the shortest way. In doing so, they bumped into novel objects that were placed en route, and jumped over phantom obstacles that were just removed previously from the familiar path. This indicated that they used proprioceptive memory rather than being guided by telereceptors (dead reckoning; [43]). Examination of speed can, thus, be highly informative about the momentary sensorimotor algorithms used by the animal.

The speed profile of the excursion undergoes a typical change in the course of Exploratory Behavior: initially, the outward part is slow and intermittent while the inward is fast and ballistic, without stopping [63], as if the home base exerts an attraction on the rat. With increasing exposure, and particularly in the course of several sessions, the proximal portion of the excursions becomes repulsive (high velocity out, low velocity in). It retains its attractive character (low velocity out, high velocity in) only at the distal portion of the excursion [65]. These properties were used to calculate a measure of familiarity for each place in the environment, and were simulated by a simple dynamical model [66].

4. Illustrations of the capabilities of SEE

4.1. Visualizing and quantifying exploratory behavior with SEE

Mathematica having been launched, and the notebook that contains the definitions of the SEE functions evaluated, one is in a position to ask the system to perform calculations and display graphics. As a first step. the files pertaining to a specific session are loaded: the smoothed time series of positions and velocities, the segmentation of the data into episodes of motion, and the categorization of the segments
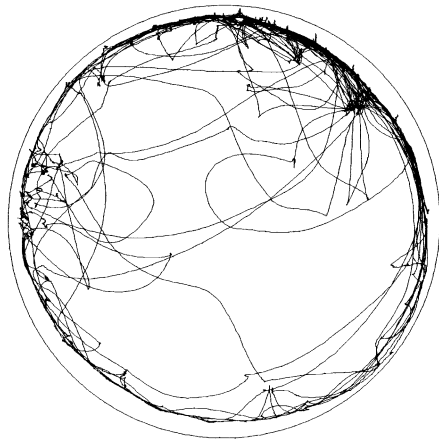
Fig. 2. The path of the rat in the arena during its fifth half-hour session. To improve visibility, arena walls are slightly removed from the path traces.

into lingering, mid speed and high speed segments (see Appendix 1.3). All the SEE functions now refer implicitly to these data. We can start looking at the animal's behavior by evaluating the command See[Pathplot[{1, Length[-data]}]] and the system will respond by displaying the graph shown in Fig. 2. This graph provides some qualitative information on the way this animal, a juvenile hooded rat, 50 days old, exploring for half an hour a 6.5 m diameter arena [65] moves during the session. The graph highlights the fact that the rat mostly moved on the periphery of the arena, progressing along relatively smooth paths.

Each of the highlighted features suggests a corresponding quantitative measure that might turn out to be useful by revealing regularity. Currently used open field measures already include the overall amount of activity, the ratio of activity spent at the periphery versus the center of the arena, activity level per fixed time intervals across the session, and the path's degree of meandering [56]. We also can perform these

measurements with our built-in functions. For example, Activity[{1, Length[data]}] 455.125 m.

We can also get a finer grasp of the dynamics of activity by computing it for successive slices of, say, five-minute intervals, thus portraying the process of occupancy of the arena: See[Map[PathPlot, TimeSlice[data, 300]]].

Here, we already get some qualitative insight: occupancy does not proceed gradually across the whole session; there is, indeed, a gradual increase in the area covered by the rat across the first three 5 min slices, but then there is an outburst of activity. It is coupled to increased space occupancy and intense exploration of the arena's center.

Fig. 3 is based on an arbitrary slicing of the path. To obtain a representation based on intrinsic constraints, it is first necessary to cleave the path into segments. This is achieved by first splicing the animal's trajectory into progression and lingering episodes. We will return to the path growth issue after visualizing lingering—the complement of progression.

The segmentation of the animal's trajectory into modes of motion (see background section) allows us to visualize the locations of lingering episodes (Fig. 4a) See[Pointplot [StopIndex]].

This graph highlights regions of higher concentration of lingering episodes, including the home base (on the right of the 12 o'clock direction). To highlight the low spatial spread of individual episodes we represent each episode by a line connecting its beginning and end, rather than by a point representing the location at which it started (Fig. 4b): See[PairPlot[lingering]].

As illustrated, in this particular session the rat did not get very far in the course of any of its lingering episodes (*M*: $20 \pm SD$ cm; [16]). TimePlot returns the spatiotemporal distribution of lingering episodes by representing the dwell time of each episode by a circle centered on the beginning of its respective location (Fig. 4c) See[TimePlot [lingering]].
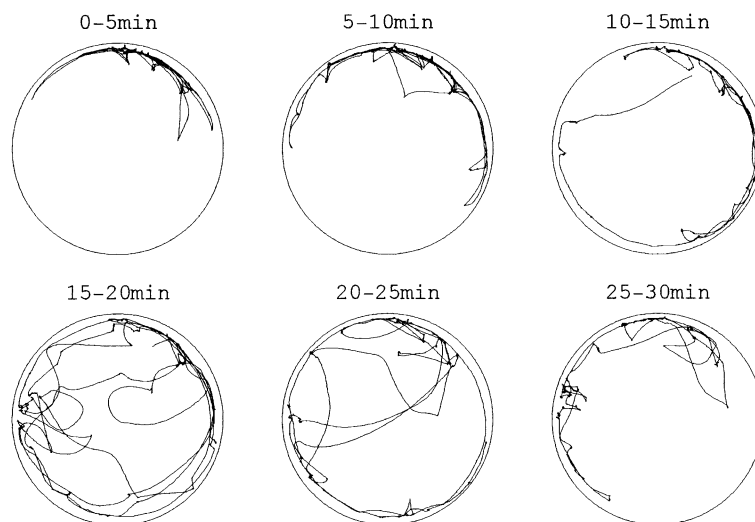

Fig. 3. The path illustrated in Fig. 1, partitioned into six successive 5 min slices.

Fig. 4. a: A representation of all the lingering episodes occurring in the arena within the same half-hour session illustrated above. Each lingering episode is drawn in the starting location of that episode. Note the higher density of lingering episodes near the home base (between 12 and 1 o'clock). b: The lingering episodes depicted in Fig. 3a are represented by a segment connecting the starting and end locations of each of them. c: The lingering episodes depicted in Fig. 3a and b are each represented by a circle whose area corresponds to the duration spent in it (dwell time) relative to the overall time spent lingering within that session. Here, the home base is clearly distinguished as a locus of frequent lingering episodes of long duration.

A more dynamical representation of the process of occupancy in terms lingering is obtained in Fig. 5, by using the command: See[TimePlot, TimeSlice[lingering, 300]].

This highlights the initial priority of the home base and the subsequent explosive increase in the spatial scatter (spread of bubbles in the arena) and temporal scatter (more evenly spread diameters of bubbles) of lingering. A numerical expression of this scatter is obtained by a measure of diversity developed in our lab: it represents the average distance between every two episodes within a given interval, weighted by the duration distribution (for principles of measurement of diversity see Ref. [54]. Diversity is defined as:

$$\sum_{i,j=1}^{n} p_i p_j d_{i,j}$$

where p_i is the proportion of time spent at stopping place i,

and $d_{i,j}$ is the distance between the location of the i-th and j-th episodes. The summation is, thus, over every possible pair of stops within an interval [64]. This measure is designed to increase as a greater distance is 'covered', and as the duration of stopping is distributed more evenly within that distance (the product of the time proportions is obviously the largest when p_i and p_j are equal).

One can easily implement this measure as a Mathematica function and add it to the set of tools (see Appendix, 1.4)

The diversity numbers corresponding to the bubble graphs displayed in Fig. 5 would be produced by evaluating Map[Diversity, TimeSlice[lingering, 300]] {94.0723, 131.645, 193.754, 339.331, 375.662, 273.02}.

As expected upon viewing the dynamic visualizations in Fig. 5, the fourth and fifth slabs show an abrupt rise in diversity.

This measure has been originally devised to capture the developmental transition from restrained to relatively free
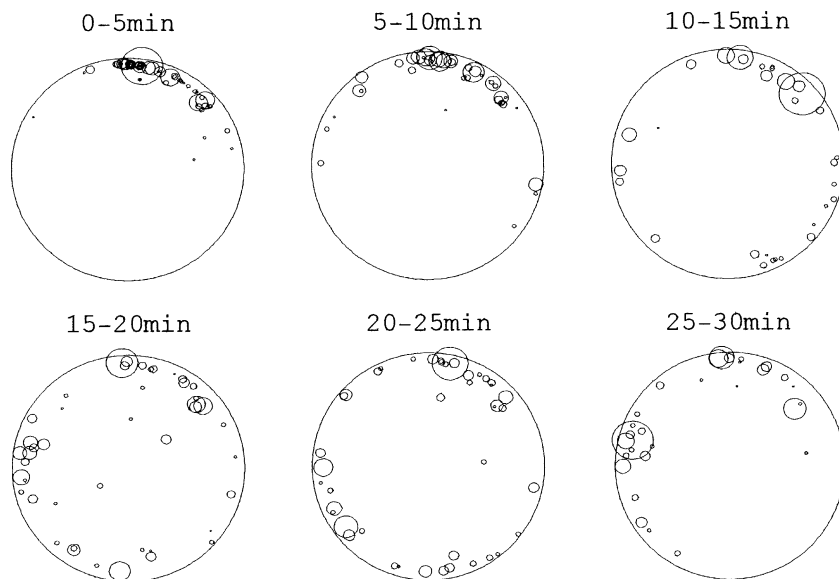

Fig. 5. A dynamic representation of lingering episodes with their dwell time. Note that the increase in activity in the first four slices is associated with an increase in Diversity (see text) in the first four slices. Notice however that activity increase from slice 2 to slice 3, whereas diversity decreases, due to a single long lingering event at 2 o'clock in slice 3.
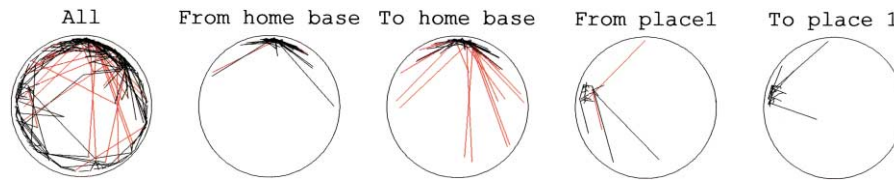
Fig. 6. A schematic representation of the movement segments. A line connecting the starting point to the ending point of a segment represents that movement segment. The first graph on the left represents all the segments in the session. The other graphs represent subsets of this, as specified. The line's color represents the maximal speed achieved in that segment: segments whose maximal speed was higher than 190 cm/s are colored in red. This speed is the lower limit of the upper quartile of the maximal speed's population. The graphs highlight the high connectivity of the home base, to which many long distance high-speed segments converge and from which many segments start. As illustrated, departures from the home base involve shorter distances and lower speeds, whereas high-speed long distance segments characterize arrivals. This is not the case with the next most visited place represented in the rightmost two graphs.

behavior. It is an instructive measure as it reflects the animal's freedom of movement in the arena [64].

The visualization suggests the hypothesis that periods of intense exploration could be better characterized by an increase in diversity rather than by a mere increase in mileage. To examine this hypothesis, all one has to do is apply the same function to a set of sessions of all the animals in the experimental population.

The function PairPlot, used before to highlight the spatial spread of lingering episodes (Fig. 4b) can also be used to highlight the connectivity between places. Fig. 6 presents five graphs generated by SpeedPairPlot, a variant of Pair-Plot, here distinguishing between high (in red), and lower speed (in black) movement segments. In the leftmost graph, the command See[SpeedPairPlot[movesegments]] shows all the segments of motion confounded.

As illustrated, the long-distance, faster progression segments tend to converge (and diverge) to particular neighborhoods. To examine the connectivity of these neighborhoods in relative isolation, we use the SEE operators that characterize segments by their ending or starting neighborhood: From[{x, y}, epsilon]: selects the progression segments whose starting point is at a distance less than epsilon from {x, y} (which will typically be the center of one of the preferred places). To[{x, y}, epsilon]: selects the progression segments whose endpoint is at a distance less than epsilon from {x, y}.

This allows us to form a preliminary opinion as to whether specific locations constitute operational places in the behavioral space of the animal.

Since the rat's home base stands out as a relatively stable structure (see Figs. 1 and 4a–c), it is only reasonable to examine the departures from and arrivals to its neighborhood. The following commands cull out the segments starting or ending at the home base's neighborhood, here designated by the coordinates {10, 350}. See[SpeedPairPlot[From[{10, 350}, 50]]] generates the departures from (b), and See[SpeedPairPlot[To[{10, 350}, 50]]] generates the arrivals to the home base (c). Note the asymmetry between arrivals and departures (Fig. 6).

The arrivals' segments reach higher maxima of distance and velocity (To home base) than the departures' segments (From home base). In other words, the rat returned to the home base non-stop from further away and at higher maximal velocities than when he left it. A similar examination of the next most preferred place (two graphs on right) shows that the arrivals/departures asymmetry does not hold for this place. PairPlot and SpeedPairPlot reveal the location of the home base by highlighting its high connectivity to the environment.

Having used lingering episodes to highlight the animal's freedom of movement in the arena (as reflected by connectivity and diversity) we can now return to progression episodes. The gradual occupancy of the arena, movement segment by movement segment, can be visualized with AnimSeg. This is a graphic function generating an animation of the animal's path in terms of the 'current' movement segment, colored from yellow to red, on the background of the whole previously traced path in gray.

(For an animated version of this type of graph see http://www.tau.ac.il/~ilan99/) A property that stands out is the gradual increase (and decrease towards the end) in movement segment lengths across the session (Fig. 7).

One feature of SEE is the constant availability of all the data within the database and the feasibility of using a single command for producing a synoptic representation of the dynamics of a particular measure across the session(s). To obtain the dynamics of movement segment lengths for all of
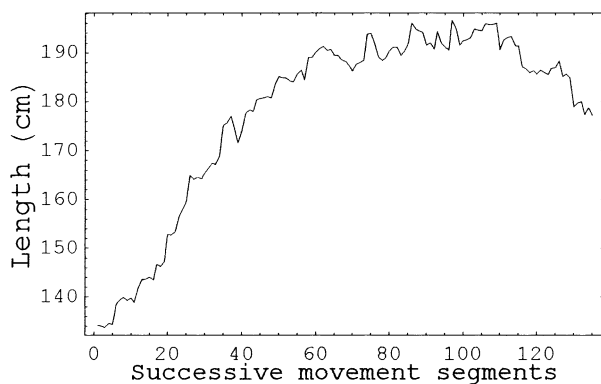

Fig. 7. A graph of the successive lengths of movement segments performed across the fifth half-hour session, smoothed by applying a moving average with a window of 40 segments.
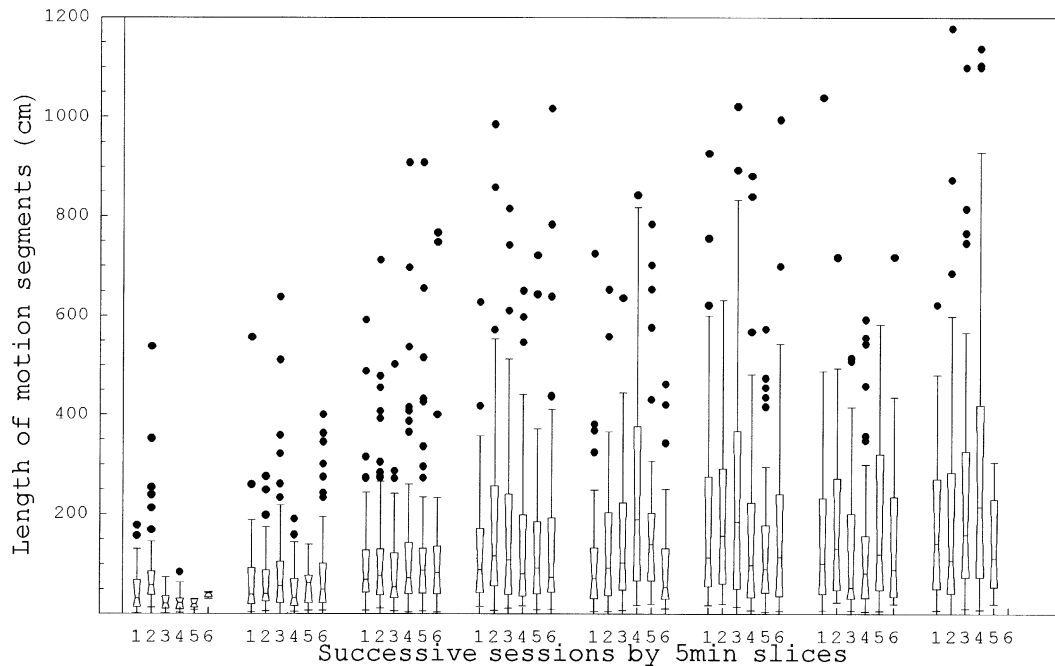
Fig. 8. A boxplot summary of the dynamics of the length of movement segments for all the sessions of a given rat. Each box represents the observations within a five-minute slice in a given session. Six boxes, thus, represent one session; the sessions are ordered from left to right, from earliest to latest. The range between the bottom and top of each box contains the central half of the data; each box is cut by a line whose height is that of the median of the data; two whiskers extend from the top and the bottom of each box to the farthest observations that are still no more than two box lengths away from the sides of the box. Outliers are shown as individual data points. Note the increase in length of movement segments both within and across sessions.

this rat's sessions we thus use the following command, requesting BoxPlot graphs for the eight successive sessions, each daily session represented by six 5 min time slices (Fig. 8). The function Script loads in turn each of the session files, performs the specified operations on it, and unloads it. BoxPlot[Script[Map[Activity, TimeSlice[movesegments, 300]], FileNames['hooded99_1_*']]]].

As shown, the increase in segment length is evident in the values of the medians, in the values of the upper quartiles, in the value of the upper whiskers, and in the values of the outliers. In other words, there is an overall increase in the range of magnitudes of individual progression segments, both within and across sessions. Some of the questions that this observation raises are how general is this build-up in amplitude of individual movement segments? Does it reflect increased familiarity with the specific arena? increased habituation to the experimental procedure? An endogenous property of innate patterns, also documented in other aspects of exploratory behavior [31,33,42,65]? Finally, is a similar growth pattern also expressed at the level of excursions in this rat?

We are, thus, led to partition the half-hour paths into excursions that start and end at the home base [32,63,65]. Two such successive excursions, bounded by stops at the home base, are illustrated in Fig. 9 top row. They constitute the 30th and 31st excursions in the rat's fifth session, and they show some of the structure unfolded during the period in which there was an outburst of activity. We present two

versions of representations of these excursions. In the first, the current excursion path, proceeding from yellow to red, is traced on the gray background of the rat's previous path history (in that session). Circles whose diameters designate the respective dwell times represent the stops performed in that excursion. The limited number of circles per excursion illustrates the upper bound on the number of stops phenomenon summarized in the background section [32].

In the second version, shown in Fig. 9 bottom row, the same two excursions are represented as a sequence of motion segments where each segment's type is designated by its color: lingering by red, mid-speed segments by green, and high-speed segments by blue (the segments are by definition bounded by stops). This colored version of PathPlot, which is generated by the function TypePathPlot is instructive because it is difficult, if not impossible, to classify segments by discerning their speed, either by observation in real time or by slowing down the video. One's power of visual discrimination of these building blocks is greatly increased by first observing the graph and then the movie of the same slab of behavior, in normal speed.

We expect this type of graph to highlight differences in the sequencing of segments.

A full dynamic representation of the rat's velocity in reference to its momentary location is provided by Phase-Plot. This function allows one to actually visualize in 3-d the animal's trajectory (Fig. 10).

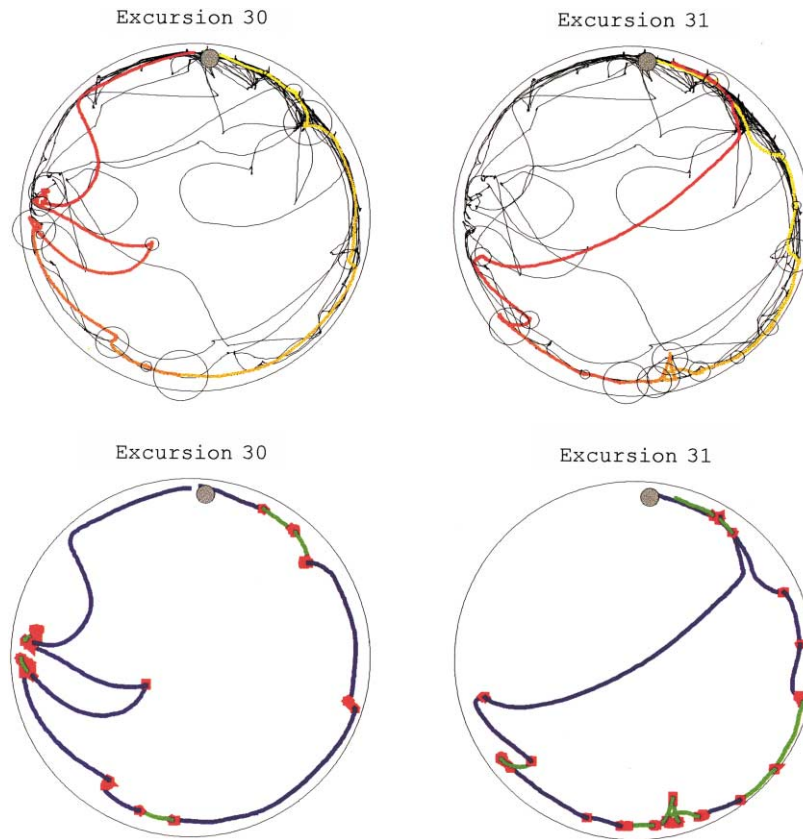This type of graph is useful for studying the relationship

Fig. 9. Top row: A representation of the path of the 30th and 31st excursions in the session. The time course is coded by the color (yellow = start, red = end); superimposed on the path are bubbles representing lingering episodes whose duration is represented by the bubble's diameter. The path of past excursions is displayed in light gray, permitting a direct perception of the build up of space occupancy from excursion to excursion. Bottom row: the same 30th and 31st excursions are represented as a sequence of movement segments where each segment's type is coded by its color: lingering by red, mid-speed segments by green, and high-speed segments by blue.

between speed, location, and curvature of the animal's path. For example, in both excursions the final movement segment towards the home base is fast and ballistic; and crossing of the center involves high speed and no stopping.

Fig. 11 provides an overview of the whole system's dynamics presented as a sequence of excursions. Several features stand out.

1. There is an overall growth in excursion length.

2. Growth proceeds first along the walls (excursions 1–25) only then into the center (excursion 26 and on).
3. Growth proceeds non-monotonically (an excursion can be shorter than its predecessor).
4. At a certain point (excursion 26) there is an abrupt burst of activity involving both the performance of full clockwise and counter-clockwise circles.

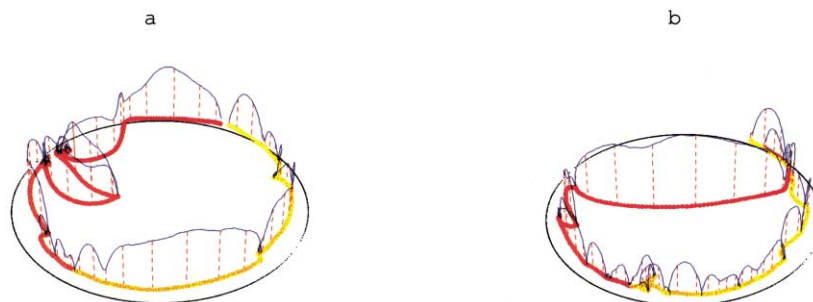The overview also highlights other dynamic aspects of the

Fig. 10. A phase space representation of the 30th and 31st excursions. The time course is coded by the color (yellow = start, red = end), and velocity is represented in blue in the third dimension. The red dashed lines show the projection of some values of velocity on the plane of the path for ease of interpretation. Both of the excursions are performed counterclockwise.
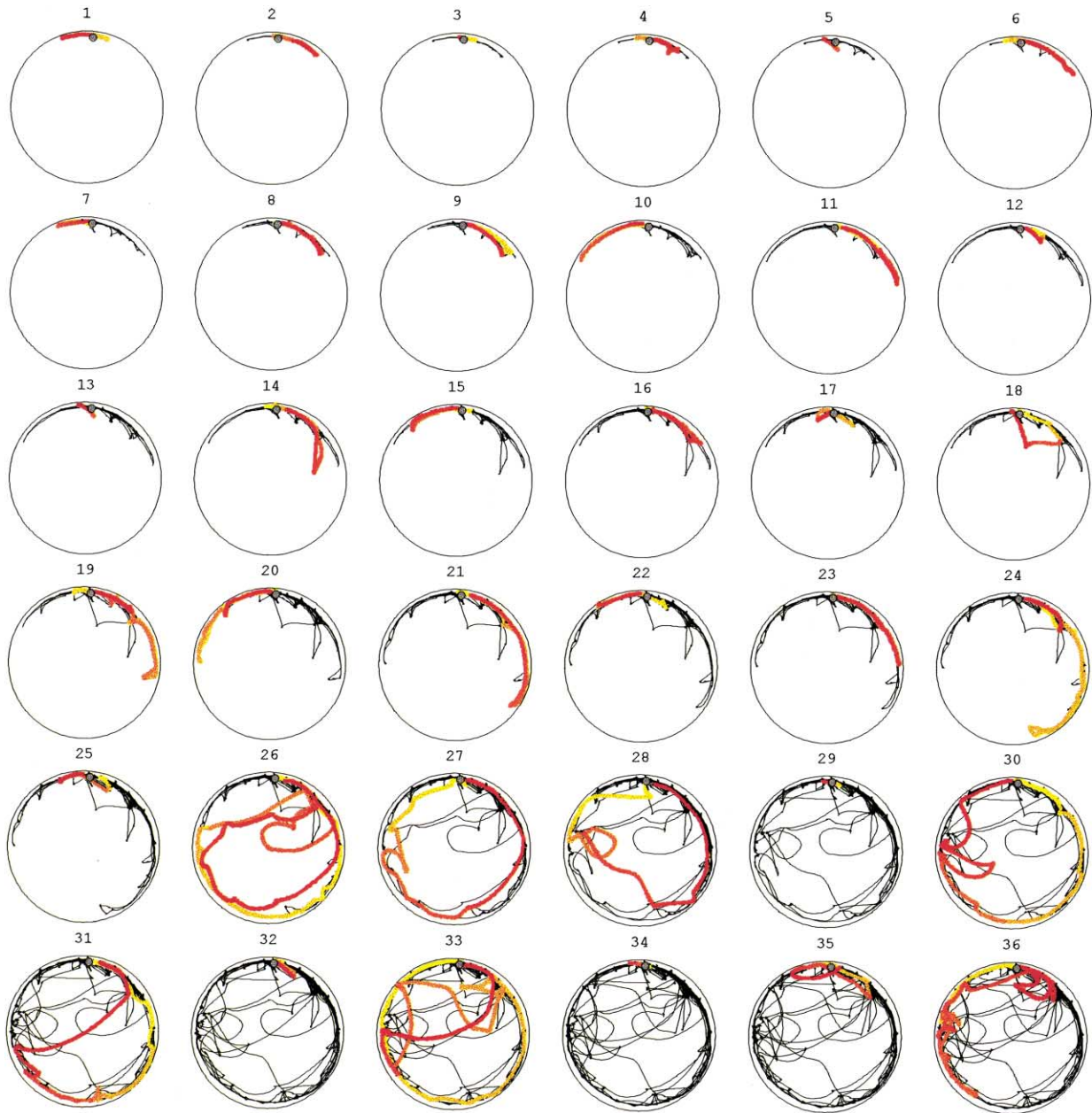
Fig. 11. A representation of all the excursions performed in session 5. The excursions are temporally ordered from left to right and from top to bottom. Each excursion is displayed in color (from yellow = start to red = end) against a light gray background of the cumulative path back to the beginning of that session.

growth process. For example, the aborted incursion appearing in excursion 30 at 9 o'clock foretells the full-blown crossing through the center occurring in excursion 31 and later in 33. In other words, a 'bud' growing in that direction precedes the growth of the path in that new direction. Once again, such observations suggest hypotheses that can be readily tested on large sets of data, and subsequently used to refine existing models of the growth process [66].

5. Discussion

Behavior genetics, behavioral pharmacology, and the neurobiology of spatial memory rely critically on the relevance, reliability and accuracy of behavioral measurements. Tests such as the Morris Water maze, the elevated plus maze, and the radial maze provide such measurements for the examination of already formulated hypotheses such as the existence and rate of spatial learning, or the presumed level of an animal's anxiety. At the present state of knowledge concerning whole-animal unrestrained behavior, however, it is also necessary to broaden and enrich our list of hypotheses regarding the structure and functional meaning of behavior. SEE has been constructed with this aim in mind. SEE is based on a series of systematically justified constructs such as the distinction between stopping

and progressing, the home base, the excursion, inbound and outbound velocities, spatial spread, diversity, dynamics of space occupancy, number of stops per excursion, and the maximal speeds attained within episodes. Each of these constructs (and additional ones, which are to be added as new preparations are examined) is defined by a corresponding algorithm. In this way, SEE establishes a framework in which the functional meaning of structures can be examined on a more refined level, yet on a large scale, using high throughput quantities of data on a variety of species, strains and preparations.

In behavioral pharmacology, measurements are typically performed in several test situations: the elevated plus-maze [57], the light/dark test [48], the open field test [35], etc. As long as all tests yield the same ranking, interpretation of the results is relatively easy. For example, BALB/cByJ mice rank higher than C57BL/6J in anxiety related parameters in four test situations, therefore, BALB/cByJ can be taken to represent a relatively anxious strain [41]. Interpretation becomes difficult, however, when uni-dimensional ranking is not possible, based on a limited set of scores taken in several test situations [3,7,27,34,59]. Another difficulty in interpretation stems from the fact that the various test situations are neither ecologically- nor behaviorally equivalent [59]. Each of the tests involves a specific range of environmental parameters (light intensity, presence/absence of walls/cliffs, etc.) and a corresponding range of behavior patterns 'which are usually ignored in favor of a (single) relatively simple pattern' (e.g. proportion of time spent in exposed areas; [59]). In the SEE framework, these difficulties are avoided because a whole set of patterns is examined dynamically within the same range of environmental parameters. Within this framework, one can examine the moment-to-moment relationship between measures, and repeat the experiment while varying systematically an environmental or pharmacological parameter. In this way one should be able to examine the resilience of significant between-strain differences vis-à-vis an environmental change, or study the changes in a list of all the behavior patterns isolated by SEE, vis-à-vis a pharmacological challenge. In a recent study performed with SEE, BALB/cJtau ranked significantly lower than C57BL/6Jtau mice in eight endpoints that could be interpreted to signify stressful behavior (e.g. shorter duration of lingering episodes, lower maximal speed of lingering episodes, shorter spatial spread of lingering episodes, lower activity at onset and during first half of session, lower diversity during first half of session; [5,17]). The lower lingering values appear to reflect the BALB/cJtau mice's paucity of scans, both in number and type, during stops. Together with a late onset of activity, these values show lower mobility (a systematically justified construct described in Ref. [31]) which could in turn be argued to embody higher anxiety. Are the nine differences between the two mouse strains differentially sensitive to a change in the lighting conditions of the arena or to conventional and novel anxiolytic agents? Can the administration

of an anxiolytic drug to the C57BL/6Jtau mice turn them into BALB/cJtau-like mice? By manipulating the environmental and the pharmacological parameters it might become possible to isolate each of the input key variables that are actually controlled by the animals during presumed states of anxiety [33].

The notion of place plays a central role in the neurosciences, in the fields of spatial memory [50,61,72,73], navigation (e.g. [24,40,67]), and the study of the hippocampus (e.g. [45,53,58]), yet there is no definition of the behavior that marks a place. In our hooded rats, lingering appears to be clustered around a small number of spatially distinct locations, including that of the home base (see Fig. 1). This suggests that lingering behavior marks places in this rat's operational world. Such a notion of place (as distinct from the notion of location, which would refer to any particular pair of coordinates in space) is obtained through the study of the geographical clustering of lingering episodes. It has the advantage of being defined precisely in terms of behavior, so that we do not have to set an a priori size for a place (as a set of coordinates circumscribed by an arbitrary radius). For instance, a place situated at the home base could be spatially extended much more than, say, a less familiar place. The validity of a place should be further supported by its connectivity to other places via movement-segments, or the performance of incoming ballistic movement-segments from relatively long distances (see Fig. 6). A preliminary impression suggests that unlike hooded rats who establish, in addition to the home base, several principal places [64], mice lingering episodes are much less, if at all, clustered geographically. Once we have at hand a geographical definition of places, it would be interesting to monitor simultaneously a rat's location and place-cells activity in the hippocampus and then examine the correspondence between behaviorally and electrophysiologically defined place fields.

As illustrated in the current paper the SEE methodology includes five steps.

1. Acquiring data, preferably at a rate of 25–30 Hz, and not lower than ~10 Hz, smoothing (Robust Lowess, [9]), computation of speeds, and segmentation of the data time-series into modes of motion [16].
2. Searching for patterns by visualization and preliminary analysis of the data on a manageable number of subjects. For example: do the animals have a home base, is there an asymmetry between arrivals to and departures from the home base, is there an increase in the length of progression segments within and/or across sessions?
3. Defining the quantity that captures the pattern. For example, the proportion of cumulative time spent at the home base, the number of visits to the home base, the slope of a regression line measuring the increase in progression segments length and excursion length, or the median maximal speed of progression and of lingering segments.
4. Computing the values of such measures on various

cross-sections of the database. For example: early activity, diversity during first half of session versus second half.

5. Assessing the dependency between already established measures. For example: what is the empirical relationship between activity and diversity, or between movement segment length and excursion length?
6. Distinguishing between true and false discoveries in the context of multiple comparisons [4].

In a previous study, we focused on the segmentation of the data into modes of motion [16]. In this study, we focused on the second step of visualization and preliminary analysis. In forthcoming studies, we treat in detail the 4th–6th steps, phenotyping the behavior of C57BL/6Jtau and Balb/cJtau mice [5,17–19]. This last study provides a list of behavioral endpoints that are compared across the two strains. After taking into account the False Discovery Rate of Benjamini and Hochberg [4], 11 out of 19 endpoints were found to differ significantly, controlling for a 0.05 level of significance. These were: early (first 5 min) activity in movement segments; overall early activity (first 5 min); lingering speed; number of stops per meter; spatial spread of lingering episodes; duration of stops; movement segment speed; early minus late diversity; early minus late activity; spatial spread of movement segments; and overall activity. One advantage of this methodology is the large number of ethologically relevant parameters obtained from a single experiment.

A prerequisite for the full applicability of SEE is the validity of the segmentation procedure [16]. So far, segmentation has been validated in several strains of rats and mice in different arena sizes (6.5 m diameter, 3.20 m diameter, 2.50 m diameter, 1.40×1.40 m, and 60×60 cm), with various tracking systems (our own custom built video tracking system, Noldus's Ethovision™, AR Cools lab custom designed tracking system, and Photobeam tracking [39], in several contexts (e.g. daylight versus total darkness conditions), and in four laboratories.

The mouse study is part of a larger scale project committed to the isolation and quantification of patterns of exploratory behavior in 17 of the most common wild type mouse strains. The software developed in the course of this process, a standard experimental protocol and endpoint tables, as well as the raw data will all be available on the web. The overall aim is to create a freely accessible web-based genetic repository of analyzed behavior together with tools for the high-throughput analysis of behavior in other strains and preparations. The collection of SEE functions will be made accessible to all researchers through the web. We anticipate that with time, old functions will be refined, and new functions will be formalized and added, especially if other researchers use SEE. In such a case, its most up-to-date version would be placed on the web, including both the old and the new functions, each ascribed to its corresponding authors. For a Mathematica 4.0 version of this paper in the form of a working manual including data,

a package with all the necessary SEE functions, and the commands that generate the visualizations presented in this paper as well as additional visualizations see the Introduction.

Appendix

A.1. Prerequisite computations and file naming conventions

The use of the visualization and quantification functions presupposes stages of data acquisition, data smoothing, and data segmentation as well as the organization of their output into a structured system of files.

A.1.1. Data acquisition

For each session, the output of the data acquisition stage should be a file consisting of the time series $\{\{t_1, x_1, y_1\}, \{t_2, x_2, y_2\}, \ldots\}$.

1. The time coordinates should be integers, expressing time units corresponding to the maximal temporal resolution of the system.
2. The coordinates should be real Cartesian coordinates (as opposed to screen coordinate), expressed in centimeters relative to a frame of reference having as origin the center of symmetry of the experimental arena.
3. The sampling rate should be at least 10 Hz. Such a sampling rate is necessary for computing meaningful velocities.

A.1.2. Data smoothing

The time series of coordinates $\{\{t_i, x_i, y_i\}\}$ is generally noisy, partly arising from the characteristics of the data acquisition device. In automatic tracking of the moving animal from a video record, for example, noise may be generated by a mechanical jitter of the video equipment, fluctuations of luminosity in the arena, and numerical rounding errors of the tracking algorithm.

We first de-noise the data by applying a robust smoothing algorithm (Robust Lowess; [9]. Once the data are smoothed, velocities can be meaningfully computed and a threshold for arrest can be determined. Such a threshold is necessary for the subsequent segmentation of the data series into modes of motion.

Given that there is a trade-off between efficiency and resolution: (i) the tracking rates we choose range between 10 and 25 frames in the European PAL video standard and

10–30 fps in the American NTSC video system; (ii) we smooth the data with a ~0.7 s to a ~1.3 s time window; and (iii) we use a 3rd or 4th polynomial to represent the raw data time series. Simultaneous use of the lowest values of these three parameters generates too noisy data, capturing only stops whose duration is 0.5 s or more. Simultaneous use of the highest values generates a nicely smoothed time series, representing a set of events, such as stops, whose duration is 0.2 s or more. This might be less necessary in the rat and more appropriate in the mouse. A systematic study of the effects of different smoothing methods and parameters on the cutoff of noise level, on the cutoff between lingering and movement segments and on other results is under way.

A.1.3. Data segmentation

A.1.3.1. Episodes of motion The input to this phase should be the smoothed data time series, whereas the output is a sequence $\{\{m_1, n_1\}, \{m_2, n_2\}, \ldots\}$ where each m_i is the index of the beginning of a motion episode, and n_i the index of its end. Determining a threshold for absolute arrest (zero velocity) and defining episodes of motion as intervals where no arrest occurs does this. For example, {102, 150} designates an episode of motion starting at the 102nd data point in the original time series, and ending at the 150th data point.

A.1.3.2. Lingering and progression segments Each episode of motion is characterized by its maximal velocity. The population of maximal velocities is then analyzed for multi-modality, and thresholds for the transition from one mode to the next can then be rigorously determined and used to characterize each of the components [16].

A.1.3.3. Gaussian model and EM algorithm The Gaussian mixture model is used for recognizing distinct components within a population. In analogy, in electrophoresis, a mixture of distinct proteins ideally yields a perfect separation of the mixture into its components, i.e. all the molecules of each component lie precisely at a distance determined by their specific mass. In practice, however, the distance a particular molecule travels is affected by, e.g. convection. As a result, the distances traveled by each type of molecule form a Gaussian. When plotting concentration against distance from origin, one gets a single curve showing peaks corresponding to the medians of each Gaussian. Given a mixture, be it of proteins or of distinct components of speed maxima in a frequency distribution curve, the number of peaks in the curve corresponds to the number of components. The actual proportion of each component can be estimated by fitting a Gaussian mixture model to the empirical curve. This model consists of a sum of distinct Gaussians weighted by their corresponding proportions. The proportions of the components in the mixture are then calculated by determining the values of the proportions that give maximum likelihood to the model.

The Expectation-Maximization (EM) algorithm [25] estimates the maximum likelihood parameters (proportions, means, and standard deviations) of a mixture with a given number of Gaussians. EM is an iterative algorithm that starts with user-given initial values, and incrementally improves the likelihood function until further iterations yield only a negligible improvement. The actual number of components of the model is determined by comparing the maximum likelihood value of an n-components mixture with that of an $(n - 1)$-component mixture until the increased number of components increases the likelihood only marginally. (For further details see Ref. [16].

A.1.4. File system and naming conventions

A good choice of naming conventions for the diverse files we need for visualization and analysis is necessary in order to produce graphs and compute endpoints for any subset of the database (e.g. all the sessions of a given animal, all the second sessions of all animals in a given project etc.). We therefore adopted the following conventions: all data files are stored in a directory C:\See data C:\Seedata, each session's data (i.e. the time series issuing from the data acquisition process) is stored as projectname_ratname_ session number.dat.

The episodes of motion are stored as projectname_ratname_ session number.mot. The velocities are stored as projectname_ ratname_session number.spd. The maximal values of the velocities within each episode of motion are stored as projectname_ratname_session number.max. The thresholds obtained in the classification process are saved as. projectname_ratname_session number.thr.

Thanks to these conventions, all the files pertaining to a given animal are easily definable. This is exploited in the implementation of the function 'Load[name]', which loads in memory all the files computed for a given animal. The Mathematica pattern matching facilities are further used in the function Script[list of computations, list of files]; this function computes a user defined list of quantities computable from the files pertaining to an animal, across a list of files. Thus: Script['{Diversity[TimeSlice [lingering, 300]], Median[SpatialSpread[movesegments]]}', FileNames ['hooded 99_*_3']] would compute the dynamics of diversity by 5 min slices and the median of the spatial spread of progression segments, for all the files corresponding to the *third* session of any animal in the project called *hooded 99*.

A.2. A formal description of SEE

The following is a collection of the data objects and functions that constitute SEE. A full practical guide to the operation of the software, as well as the implemented code is available upon request from the authors (note that one needs a working copy of Mathematica 4 in order to use the software).

A.2.1. Elementary data objects

Once a data set has been loaded through use of the Load function, all the variables refer implicitly to that data set.

data consists of a time series of the form $\{t, x, y, v\}$, where t is the time, x, y are the Cartesian coordinates and v is the absolute velocity.

contour2d represents the contour of the experimental arena viewed in two dimensions.

contour3d represents the contour of the experimental arena viewed in three dimensions (this is used to display phase plots).

motions consist of a list of couples $\{m, n,\}$. m, n are the indexes of the beginning and end of an episode of motion.

maxima consists of a list of the maxima of speed for the episodes of motion.

thresholds consists of a list $\{t_1, t_2, \ldots\}$, where t_1 is the noise level, and each subsequent t_i is the threshold separating the i-th component from the $i + 1$-th in the multi-modal analysis of the maxima.

movesegments consists of couples $\{m, n\}$, the indexes of the beginning and end of progression segments which are not of the lingering type.

lingering consists of couples $\{m, n\}$, the indexes of the beginning and end of a lingering segment.

reallingering consists of couples $\{m, n\}$, the indexes of the beginning and end of a lingering segment longer than 0.2 s.

highspeedsegments consists of couples $\{m, n\}$, the indexes of the beginning and end of a segment whose maximal value of speed is above the third threshold (if there is one).

midspeedsegments consists of couples $\{m, n\}$, the indexes of the beginning and end of a segment whose maximal value of speed is above the second threshold and below the third (if there is one).

stopindex consists of the indices of beginnings of arrests or lingering episodes.

startingpoints consists of the indices of beginnings of movesegments.

endingpoints consists of the indices of ends of movesegments.

These data objects reflect our choice of fundamental units of analysis. Different constructs defined in terms of these elementary objects will be visualized and/or computed with the help of the elementary operators.

A.2.2. Elementary operators

Besides data objects pertaining to the loaded session, SEE loads a list of definitions of operators that permit the definition of various aspects of the data. Most of them have been illustrated in the previous sections of the paper. For the sake of completeness we will list all the operators with a brief description of their function.

Distance$[\{x, y\}, \{u, v\}]$ gives the Euclidian distance between the points $\{x, y\}$ and $\{u, v\}$.

CarttoPolar$[\{x, y\}]$ translates the Cartesian coordinates $\{x, y\}$ to polar ones ({radius, angle}).

SpatialSpread$[\{m, n\}]$ gives the maximal distance between any two points belonging to the segment $\{m, n\}$.

Activity$[\{m, n\}]$ gives the total mileage in the segment $\{m, n\}$.

MaxSpeed$[\{m, n\}]$ gives the maximal speed in the segment $\{m, n\}$.

AverageSpeed$\{m, n\}]$ gives the average speed in the segment $\{m, n\}$.

Balisticity$[\{m, n\}]$ is a number between 0 and 1 computed from the oscillations of speed within the segment $\{m, n\}$. This number expresses the magnitude of changes of sign in the acceleration profile as compared to that of a maximally smooth trajectory (which has balisticity value 1 by convention). We use this as an indicator of the pre-planned nature of a segment.

From$[\{x, y\}, d]$ gives the list of segments which start from a point within a distance d from $\{x, y\}$.

To$[\{x, y\}, d]$ gives the list of segments which end at a point within a distance d from $\{x, y\}$.

StrictlyFrom$[\{x, y\}, d]$ gives the list of segments which start from a point within a distance d from $\{x, y\}$, but do not end there.

StrictlyTo$[\{x, y\}, d]$ gives the list of segments which end at a point within a distance d from $\{x, y\}$, but do not start there.

PolarFrom$[\{ro, teta\}, dro, dteta]$ gives the list of segments which start from the sector of angle 2*dteta and radii ro − dro, ro + dro, centered at the point $\{ro, teta\}$.

PolarTo$[\{ro, teta\}, dro, dteta]$ gives the list of segments which end in the sector of angle 2*dteta and radii ro − dro, ro + dro centered at the point $\{ro, teta\}$.

PolarStrictlyFrom$[\{ro, teta\}, dro, dteta]$ gives the list of segments which start from the sector of angle 2*dteta and radii ro − dro, ro + dro, centered at the point $\{ro, teta\}$, but do not end there.

PolarStrictlyTo$[\{ro, teta\}, dro, dteta]$ gives the list of segments which end in the sector of angle 2*dteta and radii ro − dro, ro + dro, centered at the point $\{ro, teta\}$, but do not start there.

Excursions$[\{x, y\}, d]$ computes the excursions relative to a home base centered at $\{x, y\}$ and of radius d. In other words, it computes the decomposition of the path of the animal into units consisting of round trips from the home base. The result is a list of lists of segments, and in each list of segments the first segment starts within the defined area of reference and the last segment ends in the same area, with no intervening stops at the area of reference.

TimeSlice$[list of segments, dt]$ partitions the list of segments into segments which occur in successive time slices of duration dt (i.e. the first sub-list consists of those segments in list of segments which occur between 0 and dt, the second of those between dt and 2*dt, etc.).

A.2.3. Graphical operators

The graphical functions of SEE compute graphical objects through the use of the elementary data objects and the elementary operators. Any of these graphical objects, or combination of them, is displayed on the screen by applying the main function See.

PathPlot[{m, n}] returns a graphics object which represents the path traced through the segment {m, n}.

PointPlot[{$m1$, $m2$, …, mk}] returns a graphics object which represents the locations of the data points $m1$, …, mk.

TimePlot[list of segments, display factor:50] returns a graphics object which represents the dwell time in each of the segments of the list by a circle. The circle is centered at the beginning of each segment, its area is proportional to the relative dwell time in the segment (relative to the sum of the dwell times over all the list of segments). The display factor controls the total area of the circles; it can be increased or decreased to make the visual display clearer, its default value is set to 50.

PairPlot[list of segments] returns a graphics object which represents the graph formed by the lines connecting the extremities of each segment in the list.

SpeedPairPlot[list of segments] returns a graphics object which represents the graph formed by the lines connecting the extremities of each segment in the list, with a color coding designating the maximal speed in each segment. The color-coding is relative to the maximal speed obtaining within all the segments (the redder the faster), or, as in Speed-PairPlot2, relative to a cutoff velocity distinguishing between the upper quartile velocities (in red) and the lower 3 quartiles (in black).

PhasePlot[{m, n}] returns a graphics object which represents the graph formed by the path between data point m and n, together with the velocity plotted in the third dimension. The time course of the path is color-coded (from yellow = beginning to red = end). The velocity is represented in azure, with red dotted lines projecting on the position plane to facilitate visualization.

AnimSeg[k, list of segments] returns a graphics object which represents the path formed by the k-th segment of the list of segments, on the background of the preceding ones. The path of the k-th segment is colored to code time (from yellow = beginning, to red = end). The paths of the segments up to the k-th are represented in gray. This function is mainly useful for creating animations of successive segments within a given list.

See[x], where x is a graphics object displays x on the screen, within the context of contour2d or contour3d. This is the function used for displaying graphics objects computed by the preceding functions, or a combination of any number of them.

References

[1] Adams LM, Geyer MA. LSD-induced alternations of locomotor patterns and exploration in rats. Psychopharmacology 1982;77:179–85.

[2] Barnett SA, Cowan PE. Activity, exploration, curiosity and fear: an ethological study. Int Sci Rev 1976;1:43–59.

[3] Belzung B, Le Pape G. Comparison of different behavioral test situations used in psychopharmacology for measurement of anxiety. Physiology and Behavior 1994;56(3):623–8.

[4] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B 1995;57(1):289–300.

[5] Benjamini Y, Drai D, Elmer G, Golani I, Kafkafi N. Controlling the false discovery rate in behavior genetics research. Behav Brain Res 2001 In Press.

[6] Berlyne DE. Conflict, arousal, and curiosity. New York: McGraw-Hill Book Company, 1960.

[7] Beuzen A, Belzung C. Link between emotional memory and anxiety states: a study by principal component analysis. Physiol Behav 1995;58:111–8.

[8] Chance MRA, Mead AP. Competition between feeding and investigation in the rat. Behaviour 1955;8:174–82.

[9] Cleveland WS. Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 1977;74:829–36.

[10] Cools AR, Brachten R, Heeren D, Willemen D, Ellenbroek B. Search after neurobiological profile of individual-specific features of Wistar rats. Brain Research Bulletin 1989;24:49–69.

[11] Cools AR, Ellenbroek BA, Gingras MA, Engbersen A, Heeren D. Differences in vulnerability and susceptibility to dexamphetamine in Nijmegen high and low responders to novelty: a dose–effect analysis of spatio-temporal programming of behavior. Psychopharmacology 1997;132:181–7.

[12] Crabbe JC, Wahlsten D, Dudek BC. Genetics of mouse behavior: interactions with laboratory environment. Science 1999;284(5420):1670–2.

[13] Crusio WE, Schwegler H, Van Abeelen JHF. Behavioral responses to novelty and structural variation of the hippocampus in mice. I. Quantitative-genetic analysis of behavior in the open-field. Behavioral Brain Research 1989;32:75–80.

[14] Delius JD. Displacement activities and arousal. Nature 1967;214:1259–60.

[15] Dempster AP, Laird NM, Rubin DB. Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion) Jour. Royal Statist. Soc. B. 1977;39:1–38.

[16] Drai D, Benjamini Y, Golani I. Statistical discrimination of natural modes of motion in rat exploratory behavior. Journal of Neuroscience methods 2000;96:119–31.

[17] Drai D, Kafkafi N, Benjamini Y, Elmer G, Golani I. Rats and mice share common ethologically relevant parameters of exploratory behavior. Behav Brain Res 2001 In press.

[18] Drai D, Kafkafi N, Benjamini Y, Elmer G, Golani I. Comparative phenotyping of exploratory behavior in two strains of mice: Balb/cJtau and C57BL/6JTAU; submitted.

[19] Drai D. An algorithmic approach to the study of exploratory behavior in rodents. Ph.D. Thesis Submitted in Tel Aviv University under the supervision of Ilan Golani; submitted.

[20] Eilam D, Golani I. Home base behavior of rats (*Rattus norvegicus*) exploring a novel environment. Behavioral Brain Research I 1989;34:199–211.

[21] Eilam D, Golani I. Home base behavior in amphetamine-treated tame

wild rats (*Rattus norvegicus*). Behavioral Brain Research 1990;36:161–70.

[22] Eilam D, Golani I. Amphetamine-induced stereotypy in rats: its morphogenesis in locale space from normal exploration. In: Cooper SJ, Hendrie CA, editors. Ethology and psychopharmacology, John Wiley & Sons Ltd, 1994. p. 241–65.

[23] Eilam D, Golani I, Szechtman H. D2 Agonist quinripole induces perseveration of routes and hyperactivity but no perseveration of movements. Brain Research 1989;460:255–67.

[24] Etienne AS, Maurer R, Berlie J, Reverdin B, Rowe T, Georgakopoulos J, Seguinot V. Navigation through vector addition. Nature 1998;396(6707):161–4.

[25] Everitt BS. Finite mixture distributions. London: Chapman and Hall, 1981.

[26] File SE. What can be learned from the effects of benzodiazepines on exploratory behavior? Neurosci Biobehav Rev 1985;9(1):45–54.

[27] File SE. Behavioral detection of anxiolytic action. In: Elliott JM, Heal DJ, Marsden CA, editors. Experimental approaches to anxiety and depression, Chichester: Wiley, 1992. p. 25–44.

[28] Gallistell CR, Krammer AE. Computations on metric maps in mammals: getting oriented and choosing a multidestination route. J Exp Biol 1996:211–7.

[29] Gerlai R. Molecular genetic analysis of mammalian behavior and brain processes: caveats and perspectives. The Neurosciences 1996;8:153–61.

[30] Gingras MA, Cools AR. Different behavioral effects of daily or intermittent dexamphetamine administration in Nijmegen high and low responders Psychopharmacology 1997;132:188–94.

[31] Golani I. A mobility gradient in the organization of vertebrate movement: the perception of movement through symbolic language. The Behavioral and Brain Sciences 1992;15(2):249–66.

[32] Golani I, Benjamini Y, Eilam D. Stopping behavior: constraints on exploration in rats (*Rattus norvegicus*). Behavioral Brain Research 1993;53:21–33.

[33] Golani I, Kafkafi N, Drai D. Phenotyping stereotypic behavior: collective variables, range of variation and predictability. Applied Animal Behavior Science 1999;65:191–220.

[34] Griebel G, Belzung C, Perrault G, Sanger DJ. Differences in anxiety-related behaviours and in sensitivity to diazepam in inbred and outbred strains of mice. Psychopharmacology 2000;148:164–70.

[35] Hall CS. Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. J Comp Psychol 1934;18:385–403.

[36] Hoffmann G. Experimentelle und theoretische analyse eines adaptiven Orienterungsverhaltens: die optimale Suche der Wuste-nassel *Hemilepistus reaumuri*, Audouin und Savigny(Crustacea, Isopoda, Oniscoidea) nach ihrer Hohle. PhD thesis, Regensburg 1978.

[37] Hoffmann G. The random elements in the systematic search behavior of the desert isopode *Hemilepistus reaumuri*. Behav Ecol Sociobiol 1983;13:81–92.

[38] Hoffmann G. Optimization of Brownian search strategies. Biol Cybern 1983;49:21–31.

[39] Kafkafi N, Mayo C, Drai D, Golani I, Elmer G. SEE analysis of photo-beam raw data. J Neurosc Methods 2001 In Press.

[40] Knierim JJ, Kudrimoti HS, McNaughton BL. Interactions between idiothetic cues and external landmarks in the control of place cells and head direction cells. J Neurophysiol 1998;80(1):425–46.

[41] Lepicard EM, Venault P, Perez-Diaz F, Joubert C, Berthoz A, Chapouthier G. Balance control and posture differences in the anxious BALB/cByJ mice compared to the non anxious C57BL/6J mice. Behav Brain Res 2000;117:185–95.

[42] Lorenz KZ. Uber die Bildung des Instinktbegriffes. Naturwissen-schaften 1937;25:289–331.

[43] Lorenz KZ. King Solomon's ring. London: Methuen, 1952.

[44] Lorenz KZ. The foundations of ethology. Springer, 1981.

[45] McNaughton BL, Barnes CA, Gerrard JL, Gothard K, Jung MW, Knierim JJ, Kudrimoti H, Qin Y, Skaggs WE, Suster M, Weaver KL. Deciphering the hippocampal polyglot: the hippocampus as a path integration system. Journal of Experimental Biology 1996;199(1):173–85.

[46] Menzel ME. Cognitive mapping in chimpanzees. In: Hulse SH, Fowler H, Honig WK, editors. Cognitive processes in animal behavior, 1978. p. 375–422.

[47] Merimee P. Carmen and Other Stories. Oxford: World's Classics, 1999.

[48] Misslin R, Belzung C, Cogel E. Behavioral validation of a light/dark choice procedure for testing anti-anxiety agents. Behav Proc 1989;8:119–32.

[49] Mittelstaedt H. Analytical cybernetics of spider navigation. In: Barth FG, editor. Neurobiology of arachnids, Berlin, New York: Springer, 1985. p. 298–316.

[50] Morris R. Developments of a water-maze procedure for studying spatial learning in the rat. J Neurosci Meth 1984;11:47–60.

[51] Muller M, Wehner R. The hidden spiral: systematic search and path integration in desert ants, *Cataglyphis fortis*. J Comp Physiol (A) 1994;175:525–30.

[52] Nolan PM, Kapfhamer D, Bucan M. Random mutagenesis screen for dominant behavioral mutations in mice. Methods 1997;13:379–95.

[53] O'Keefe J, Burgess N. Geometric determinants of the place fields of hippocampal neurons. Nature 1996;381(6581):428–.

[54] Patil GP, Taillie C. Diversity as a concept and its measurement. J Am Statis Ass 1982;77:548–60.

[55] Paulus MP, Geyer AG. Environment and unconditioned motor behavior: Influences of drugs and environmental geometry on behavioral organization in rats. Psychobiology 1997;25(4):327–37.

[56] Paulus MP, Dulawa SC, Ralph RJ, Geyer MA. Behavioral organiza-tion is independent of locomotor activity in 129 and C57BL/6JTAU mouse strains. Brain Res 1999;835(1):27–36.

[57] Pellow S, Chopin P, File SE, Briley M. Validation of open:closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. J Neurosci Methods 1985;14:149–67.

[58] Poucet B, Benhamou S. The neuropsychology of spatial cognition in the rat. Crit Rev Neurobiol 1997;11(2–3):101–20.

[59] Rodgers RJ. Animal models of 'anxiety': where next? Behavioural Pharmacology 1997;8:477–96.

[60] Rodgers RJ, Dalvi A. Anxiety, defence and the elevated plus-maze. Neurosci Biobehav Rev 1997;21(6):801–10.

[61] Silva AJ, Giese KP, Fedorov NB, Frankland PW, Kogan JH. Molecular, cellular, and neuroanatomical substrates of place learning. Neurobiology of Learning and Memory; 1998;70(1–2):44.

[62] Sinnamon HM, Karvosky ME, Ilch CP. Locomotion and head scanning initiated by hypothalamic stimulation are inversely related. Behav Brain Res 1999;99(2):219–29.

[63] Tchernichovski O, Golani I. A phase plane representation of rat exploratory behavior. Journal of Neuroscience Methods 1995;62: 21–27.

[64] Tchernichovski O, Benjamini Y, Golani I. Constraints and the emergence of freedom in the ontogeny of rat exploratory behavior. Behaviour 1996;133:519–39.

[65] Tchernichovski O, Benjamini Y, Golani I. The dynamics of long-term exploration in the rat. Part I. A phase-plane analysis of the relation-ship between location and velocity. Biol Cybern 1998;78:423–32.

[66] Tchernichovski O, Benjamini Y. The dynamics of long term exploration in the rat—Part II. An analytical model of the kinematic structure of rat exploratory behavior. Biol Cybern 1998;78(6):433–40.

[67] Thinus-Blanc C. Animal spatial cognition: behavioural and brain approach. Singapore: World Scientific, 1996.

[69] Van de Rijt-Plooij HH, Plooij FX. Growing independence, conflict and learning in mother–infant relations in free-ranging chimpanzees. Behaviour 1987;101(1–2):1–86.

[70] Wehner R, Srinivasan V. Searching behavior of desrt ants, genus *Cataglyphis* (Formicidae, Hymenoptera). J Comp Physiol (A) 1981;142:315–38.

[71] Whishaw IQ, Cassel JC, Majchrzak M, Cassel S, Will B. 'Short-stops' in rats with fimbria–fornix lesions: evidence for change in the mobility gradient. Hippocampus 1994;4(5):577–82.

[72] Whishaw IQ, Mittleman G. Visits to starts, routes and places by rats (*Rattus norvegicus*) in swimming and navigation tasks. J Comp Psychol 1986;100:422–31.

[73] Whishaw IQ. Place learning in hippocampal rats and the path integration hypothesis. Neurosci Biobehav 1998;22(2):209–20.

[74] Wolfram S. The Mathematica Book. 3rd ed.. Wolfram Media/Cambridge University Press, 1996.