8-19-2013

Multi Scale Interactions in Biological Motion Perception

Zsolt Palatinus
University of Connecticut - Storrs, zsolt.palatinus@uconn.edu

Recommended Citation

Palatinus, Zsolt, "Multi Scale Interactions in Biological Motion Perception" (2013). *Doctoral Dissertations*. 228.
http://digitalcommons.uconn.edu/dissertations/228

Multi Scale Interactions in Biological Motion Perception
Zsolt Palatinus, M.A.
University of Connecticut, 2013

This dissertation investigates the potential contribution of fractal fluctuations of head sway in the time evolution of visual recognition in biological motion perception. The first experiment found no difference in recognition times when point light display (PLD) activities are shown either from a fixed or a moving point of observation. The second experiment, using head tracking, multifractal analyses, and geometrical manipulations in the PLDs found that (1) the multi-scale fractality of head sway is different before and after recognition, and (2) the time-evolution of the multifractal spectra predicts recognition. The third experiment manipulated both the geometrical qualities of PLDs and the context of the presentation (e.g., showing the PLDs multiple times, priming the PLDs with congruent or incongruent images of objects, and adding contextual changes). In all cases, the time-varying multifractal structure of head sway predicted changes in visual perception, in particular, the transition from exploration to recognition. Event history analyses indicated a reliable contribution of the width of the multifractal spectrum width to the evolution of recognition. Despite a wide range of geometrical and contextual manipulations on the PLDs, multi-scale interactions remained a strong index of the self-organization of biological motion perception.

Multi-Scale Interactions in Biological Motion Perception

Zsolt Palatinus

M.A., University of Szeged, 2007

A Dissertation Submitted
in Partial Fulfillment of the
Requirements of the Degree of
Doctor of Philosophy
at the
University of Connecticut

2013

APPROVAL PAGE

Doctor of Philosophy Dissertation

Multi Scale Interactions in Biological Motion Perception

Presented by Zsolt Palatinus, M.A.

Major Advisor____

James A. Dixon

Associate Advisor____

Claudia Carello

Associate Advisor____

Claire F. Michaels

Associate Advisor____

Jeffrey-Kinsella Shaw

University of Connecticut

2013

Information is shaped by the mutuality of perceiver and environment

Emission theory (Empedocles, Plato, Euclid, Ptolemy) proposed that visual perception is accomplished by rays of light emitted by the eyes. Intromission theory (Aristotle, Galen, Alhazen), in contrast, posited that visual perception comes from something representative of the object entering the eyes. These competing ancient ideas represent two ways of trying to understand vision. The former approach seeks an explanation in looking at processes generated by the organism, the latter suggests that visual perception is best explained by processes coming from the environment. A third perspective, pursuing an explanation in the interaction of processes generated by both sides was recognized early on in the history of studying perception (Reid, 1785; Mill, 1843; James, 1890) but it was taken into serious consideration by Gibson (1966, 1979). The interactive approach is grounded in the realization that there is very complex and rich structure in both the electromagnetic radiation reflected from surfaces and within the organism acting in order to detect meaningful and reliable patterns in the radiation. Rather than attributing the essence of visual perception to either side, this third approach seeks explanations in terms of the interaction of the complex patterns in the Organism-Environment system (Dixon, Holden, Mirman & Stephen, 2012; Stephen, Anastas & Dixon, 2012).

In *ecological optics* (Gibson, 1979), information is available in the optical distribution but it has to be attended and it can be clarified to various levels of detail. Attention and clarification are dependent on the observer's intentions, capacities and behavior. What matters to the animal is not *inferred* from summing up local intensities, tracing contours or point trajectories, rather, it *emerges* from an unbroken chain of organism-environment interactions (Reed, 1996; Turvey, Shaw, Reed, & Mace, 1981). Visual perception and perception in general are ways in which living systems situate themselves with respect to the actualities of their surroundings. Albeit a rather special type of connection to the world, visual perception is interwoven within all other processes in which the organism continues to exchange matter and energy with its environment. This dissertation seeks to establish explicit connections between visually guided behavior and the multi-scale flux of energy patterns around and within the perceiver.

The Gottliebian framework

Conceptual support for vision being interactive and nested within all other life processes comes from Gilbert Gottlieb's (1998, 2000) theory of probabilistic epigenesis. This theory describes the development of structures in a living system in terms of bidirectional relationships among four large scales of organization (genetic, neural, behavioral and environmental, see Figure 1). In the explanation of structural changes, the theory of probabilistic epigenesis shifts the focus from distinguished scales or localities to their inter-level relations, from traditional causal explanations to the relationships between levels of organization both within the organism and at the borders of exchange with the environment (Valsiner, 2007). The parallel to be stressed here is that Gottlieb's theory denies that higher level behavioral and somatic traits can be traced back to either genes or nurture; it also implies that visual awareness *is not a consequence* of “processing in the brain” or “some geometrical property of the environment”. Gottlieb's model deals with a much larger class of phenomena, the overall unfolding of living systems. Visual perception can be considered as a subset of the original model, nested within the totality of life-sustaining processes. This is not an extension or an addition to the original model; rather, it is an expansion that shifts the focus to the interacting levels that may play a role in the self-organization of visual awareness. Figure 2 depicts a subset that includes four levels of organization representing the perceptual system of an organism surrounded by ambient optic flow (posture, head, eye and light). As in the original model, the arrows represent the flow of influences among the levels. The levels are characterized by fluctuations that take place in the perceptual system and in the optical distribution. Bi-directionality entails that all levels are reacting to and forcing all other levels above and below. Thus, fluctuations on any given level reflect multi-scale interactions among levels, participating parts and processes. When a rich enough measurement is taken somewhere in this system, the fluctuations of the values bear marks of the changes in participating levels. Depending on the frequency and length of observation, some interrelations may be captured and tested for their contribution to other relations. This is a massively inter-dependent system of impredicative relations in which salient, observable behaviors can often be the result of various non-obvious factors distributed throughout the contexts of the participating processes. Among the concepts that make this theory well-suited for the ecological approach is the emphasis on rejecting the idea of privileged scales. Any event within the organism-environment system is contingent upon both small constituent events acting “from below” and larger contextual events acting “from above”. Moreover, the interpretation of the organismic and environmental processes as connected by cascades of energy dissipating through the system offers alternative ways of studying perceptual and cognitive behavior. The foundation for this inquiry is the hypothesis that the metabolism of energy gradients patterned by structures and processes is the supportive basis for life and behavior, including perception and cognition (Turvey, Shaw, Reed, & Mace, 1981). The fluctuations, inhomogeneities of measured intensities, always reflect some sort of energy-matter exchange regardless of whether they originated in the environment or within the organism. It is the seamless translation of diffusive energy flow from “outside-in” and “inside-out” that feeds self-organizing non-linear processes. The last couple of decades in various fields of science concluded that the fluctuation of biological processes are often best characterized as fractals. For example, heartbeats (Peng, Havlin, Stanley, & Goldberger, 1995), posture (Duarte & Zatsiorsky, 2001), DNA (Peng, Buldyrev et al., 1994), finger tapping (Lemoine, Torre, & Delignières, 2006) and human gait (Hausdorff et al., 1996) were shown to exhibit fractal fluctuations.

Multifractality

Fractality essentially means a lack of a characteristic scale and a nesting of constituent processes. The abundance of 1/f noise and fractal scaling in measurements taken on natural systems indicate interdependencies analogous to the ones warranted by Gottlieb's model. Since the irregularities of natural system's fluctuations are the consequence of bidirectional influences on very many simultaneously running and interacting processes, traditional dynamical analyses that aim to reduce the dynamics onto a small number or governing functions or single exponents cannot capture the complexity at hand. It has been shown repeatedly, that the complex structure of fluctuations in natural systems is best characterized by multiplicative cascade dynamics (Ivanov et al., 2001; Ihlen & Vereijken, 2010, Stephen et al., 2012). Multifractality is a consequence of multiplicative cascade dynamics. Visual perception requires structured optical distribution paired with a structured neural and muscular behavior from the organism. In recent years, a growing body of research suggest that both the structure of the optical distribution and the structure of fluctuations of the visual perceptual systems are also best characterized as multifractal. Let us review briefly the levels of Figure 2 focusing on recent findings emphasizing multifractal structure of each level.

In natural scenes the light coming from the sun or other light sources gets scattered by objects and surfaces of various size, ranging from particles much smaller than the wavelength of the light up to the size of mountains. The result of this scattering is a very rich and dense structure that “fills up” the illuminated environment to the degree that every point becomes an “intersection of rays coming from all directions” (Gibson, 1979, p.50). Scattering takes place when the electromagnetic wave (light ray) encounters an obstacle or non-homogeneity such as a surface, or a particle. The light wave perturbs the obstacle periodically by interacting with the electron orbits of its constituent molecules. The resultant oscillation is manifest as a source of electromagnetic radiation in its own right and it creates the scattered light. As has been shown recently (Hahn, 2009), scattering is much more than “bouncing off” the surface of obstacles; a rather complex interaction takes place on the atomic level between the light and the scattering particles. As a result of these interactions, scattered light is inhomogeneous in intensity over time and its fluctuations are non-Gaussian (Ding, Wang, Nguyen, Boppart, & Popescu (2008). Recent work has shown that the nature of this fluctuation is best interpreted using multiplicative cascade modeling (Shayeganfar, Jabbari-Farouji, Movahed, Jafari & Tabar, 2009). These studies strongly suggest that the “dense structure” that fills up the optical distribution is characterized by multiplicative cascades.

If we now step up to a higher level and investigate the properties of the “arrested” light in images that are captured by optical devices, we may see that natural images retain the complex multi-scale structure of the fluctuations that created them. This is evidenced by, for instance, the power-law behavior of the power spectrum (Field, 1987) that yields scale invariant two-point correlations in contrast. Various ways of approximating the multifractal structure of images has been shown to provide a more complete description of visual scenes (Nevado, Turiel & Parga, 2000).

When engaged with patterns of the ambient optical distribution or with recorded images of it, eyes are never at rest. While scanning the optical patterns, eye movements include larger and smaller saccades, slow drifts, and a high frequency tremor with an amplitude smaller than 1' (Ratliff & Riggs, 1950; Steinman, Haddad, Skavenski, & Wyman, 1973). How much these nested movement scales contribute to visual awareness has been debated by researchers ever since the richness of fixation structure was discovered. One way to study the contribution of eye movements is to control for them by stabilizing the eyes during perceptual tasks. Early research in stabilized conditions (Ditchburn & Ginsborg, 1952; Riggs & Ratliff, 1952; Yarbus, 1967) showed that over a period of several seconds, images tend to fade away. In other words, with long stimulus presentations, contrast sensitivity seems to get reduced dramatically especially in low spatial frequencies (Koenderink, 1972). With brief presentations (500 – 1000 ms), researchers found no significant effect of image stabilization on either acuity or contrast sensitivity (Keesey, 1960; Tulunay-Keesey & Jones, 1976). In visual discrimination tasks, however, as it has been shown recently (Rucci & Desbordes, 2003) percentage of correct responses are significantly lower in stabilized conditions than in the presence of natural eye movements even when the stimulus was presented for 500 ms. This suggests that in visual tasks, such as discrimination or recognition, eye movements contribute to visual experiences. In other words, activity on the scale of tremor might be just as necessary to visual recognition as large scale exploratory saccades and movements of the head, trunk and body.

The nested fluctuations of smaller and larger saccades have been shown to exhibit multifractal structure in various visual tasks (Aks, Zelinsky & Sprott, 2002; Shelhammer & Joiner, 2003; Stephen, Boncoddo, et al., 2009). Moreover, vision related somatic fluctuations are not exclusive to eye movements. Gibson's (1966) suggestion to consider the senses as perceptual systems instead of as separate stimulation-specific mechanisms is now supported by a large body of research. Eye movements are nested within head sway, and head sway is nested within postural sway. Both head movements and postural sway (upright and seated) are affected by the visual environment and play their role in active exploration of the optical patterns. Both head sway and postural sway exhibit multifractal structure (Morales & Kolaczyk, 2002; Hermann, 2005; Palatinus, Kelty-Stephen & Dixon, 2013). Subtle fluctuations in either seated or standing posture and head sway may moderate the effects of optical energy arrays upon the perceptual system in various visual tasks (Paulus, Straube, & Brandt, 1984; Bronstein & Buckwell, 1997; Ehrrenfield, Guerraz, Thilo, Yardley, & Gresty, 2003), even in visual search tasks (Stephen & Mirman, 2010; Stephen & Anastas, 2011).

Fractal fluctuations seem to characterize every aspect of visual perception from the scattering of light, the patterns in the optical field, through eye movements to the movements of the body supporting those eye movements. In other words, visual awareness entails a rich dynamics of interactions among the levels of organization of the living system and its environment. The flow of energy on all participating scales is reflected by fluctuations – seemingly random changes in intensities. Depending on the resolution of the measurement taken at any site of these interactions, many of the interconnected power-law relationships are present in the data. It is important to emphasize that individual readings represent arrested states of dynamical systems. However, submitting subsequent data series to fractal analyses may open a window on transitions tied to observable changes in perception. Recent findings support the hypothesis that time-dependent changes in the overall fractal scaling may reflect transitions in perception or cognition (Ihlen & Vereijken, 2010; Dixon, Holden, Mirman & Stephen, 2012). What visual search, recognition, cognitive problem solving, remembering, way-finding and judging properties of the environment have in common is that the resolution of these perceptual or cognitive problems is preceded by exploration. Exploration for a self-selected or an instructed goal or solution has been connected to the concept of *intermittency* (Ihlen & Vereijken, 2010). Intermittency is a concept that captures the emergent change in behavior quantitatively, through the analysis of the inhomogeneously distributed variability within the recorded response series. In these analyses emergent change is often indexed by changes in the width of the estimated multifractal spectrum. These emergent changes are thought to signal wide range shifts in the nested perceptual or cognitive system that is in the business of searching or exploring patterns in order to find solutions or “meaning” in its complex environment. This suggests that the amount of change can also predict the time of qualitative changes, such as the shift from exploration to recognition.

This dissertation aims to approach the problem of visual recognition in ways that are consistent with Gottlieb's requirements for bidirectional, across-level interactions and Gibson's requirements for organism-environment mutuality. The central hypothesis of this work is that the registration of specific patterns of energy distributions manifests in multifractal, scale-invariant fluctuations of exploratory movements during the time course of visual exploration and the changes in the quantitative structure of those fluctuations contribute significantly to recognition. In order to test this hypothesis, multifractal analysis was applied to time series collected from head sway while participants tried to identify various human behaviors depicted in point-light displays. This experimental task, developed for studying biological motion perception, has several advantages for present purposes.

Biological motion perception

The term *biological motion* was introduced by Johansson (1973) to refer to visual motion patterns generated by moving terrestrial bipeds and quadrupeds. In order to demonstrate the sensitivity of human perceivers to biological motion patterns, Johansson developed a minimal information display, the *point light display* (PLD). PLDs are film, video or computerized 2D displays that reduce the recorded motion of an animal to the motion of small dots—typically indicating the major joints—against a homogeneous background. Human perceivers are relatively fast and reliable in identifying the animal being presented as well as providing correct descriptions of many aspects of the observed actions. For example, from human PLDs, observers can identify gender (Barclay et al., 1978), emotions (Dittrich et al., 1996), familiar individuals (Cutting & Kozlowski, 1977), their own motion patterns as distinct from someone else's (Beardsworth & Buckner, 1981). They can estimate the weight of lifted objects (Bingham, 1993) or whether the presented human was carrying a baby or groceries (Hodges & Lindheim, 2006). However, perceivers’ performance is limited by the number of dots (Neri et al., 1998) and the orientation of the display. Upside-down presentation yields almost no recognition (Sumi, 1984; Shipley, 2003).

Despite these limitations, perception of biological motion remains remarkably robust even in severely impoverished or ambiguous conditions. For example, very short duration presentation (Johansson, 1973), blurred dots (Ahlstrőm et al., 1997) or embedding the PLD in an array of dynamic noise dots (Bertenthal & Pinto, 1994) still allows human observers to make correct judgments. Beintema and Lappe (2002) reported reliable performance even when the dots appeared randomly along the line of the limb segments in each frame of the animation. Moreover, newborns and infants show more interest in PLDs of humans than random dot motion (Bertenthal et al., 1984; Simion et al., 2007). Other species (e.g., cats, Blake, 1993; chicks, Vallortigara et al., 2005) have been reported to perceive biological motion.

Traditionally, biological motion perception has been treated as a very complex mapping problem. A general assumption is that the visual system takes the individual moving dots of the 2D display as a starting point and integrates the dots into movement trajectories, then binds the trajectories to form body segments. The body segments finally get integrated into a single coherent percept of a moving animal. However, since there is no unique mapping from 2D dot trajectories to a 3D model, biological-motion perception theories are forced to find ways to reduce the indefinitely many solutions to the correct ones in order to keep the assumption of integration intact. Candidate solutions for this reduction reflect the historical choices: Either mental processing (organism) or invariants of projective geometry (environment) are proposed. Reducing the number of possible 3D interpretations through mental processing is usually aided by positing some kind of preexisting knowledge. The required knowledge may include anatomical constraints (Hoffman & Flinchbaugh, 1982; Webb & Aggarwal, 1982) or a set of encoded patterns (sequences of snapshots of body movements) that some neural processes compare to the visually perceived patterns (Giese & Poggio, 2003). Others tried to decompose the PLD according to a set of rules that stem from biomechanical and functional constraints that are stored in memory (Troje, 2002). Proposed solutions based on projective geometry looked for invariants in the geometry of the projections that may reduce the ambiguity. Example invariants of projective geometry that were hypothesized to reduce 3D to 2D ambiguity are the cross ratio and differential motion parallax (Cutting, 1986). A complete set of invariants that would disambiguate PLDs has not been suggested so far.

Point-light displays were chosen as visual display for testing multifractal interdependencies as predictors for recognition for a number of reasons. First, recognizing biological motion within a PLD is a phenomenon equally challenging for both mentalistic and geometrical hypotheses. Most visual “cues” that are commonly suspected by cognitivist theories to provide a basis for visual processing are absent from PLDs. There are no depth cues, motion parallax is ambiguous, binocular disparity plays no role, there is no information about contrasts, edges, textures, surface gradients, supporting surface or horizon. Since pictorial information is not present, perceivers do not recognize PLDs when looking at single frames and their robust performance on the movies requires dynamical theories of biological motion perception where besides location and trajectories, time is accounted for. A further advantage of this particular stimulus paradigm is that it is relatively easy to manipulate geometrical qualities in order to create versions of the display that may provide additional hints and evaluation points for competing theories. In this dissertation, projection related manipulations included (1) orthographic projection, (2) projection to a plane that is parallel to the supporting surface (top view), (3) scrambling the points, and (4) introducing a moving point of observation.

Manipulations at the level of projective geometry

(1) Orthographic projections are a subset of parallel projections where the projector lines intersect the projection plane at right angles. Because the projector lines are parallel there is no diminution of projected size with distance. Moreover, *differential motion parallax*, an important invariant of the optic flow is missing or rather creates ambiguity in parallel projections (Cutting, 1986, p.191). A careful comparison of polar projection vs parallel projection in biological motion perception remains to be done. Here, parallel projection was chosen to demonstrate the robustness of recognition even in the absence of the candidate invariants of projective geometry.

(2) Another perspective-related manipulation was rendering PLDs from viewing angles intersecting the supporting surface at right angles, providing a top or bottom view of the actor (parallel projection does not discriminate between top/bottom or front/back). Pilot trials showed that PLDs observed from the top view were more difficult to recognize than from any other viewing angle. Perhaps this finding is related to the inversion effect (Sumi, 2000; Pavlova & Sokolov, 2000) insofar as being less frequently experienced in natural scenes. In Experiment 2b, PLDs starting from the top view and having the camera descend to a side view were used to create situations in which recognition gets delayed to a varying extent.

(3) In order to establish a baseline at which no recognition occurs, a set of PLDs with scrambled points were created. There are different methods to break the original configuration. For example, Hirai, Kaneoke, Nakata, & Ryusukekakigi. (2008) shuffled the dots in such a way that each dot retained its own trajectory but their initial positions were randomly switched. In the present work, each dot representing a joint was displaced from the corresponding joint by a random distance between zero and a leg-length in a random direction. In the resultant PLDs, the dots still trace the movements of the corresponding joints

but not at the location of the joint. In pilot work, this manipulation made recognition nearly impossible.

(4) Traditionally, PLDs are recorded/displayed according to a fixed point of observation. Orientation (Sumi, 1984; Pavlova & Sokolov, 2000) and viewing angle with respect to the observer (Cutting & Kozlowski, 1977) have been manipulated between trials but PLDs are usually presented from an unchanging perspective. In real-life situations, however, making sense of ongoing events in the environment is hardly any more difficult from a moving point of observation. Rather, a changing perspective may reveal more invariant structure in the optic array (Runeson, 1988). For obvious reasons, animals have to be aware of the presence and actions of other creatures even while moving about their goals. For example, part of playing a team sport successfully is being able to recognize the other players' actions, intentions and what the actual situation affords even while the point of observation is being altered by the perceiver's own movements.

PLD presentations with changing perspective can be prepared by assigning the plane of projection to virtual camera movements. Obviously, this technique does not make the observer move with respect to the scene, rather, it is analogous to moving the camera in cinematography (tilting, panning, dollying or trucking). A comparison between stationary and moving points of observation in cinematography was pioneered by Hochberg (1986). In movies and animated movies various camera movements do not impair perception; rather, the technique can be used skillfully as a story-telling tool, enhancing the viewer's engagement in the scene in artistic ways (Mitry, 1990). The result is a projection that is observed from a more or less stationary observation point but the projection itself is a consequence of a changing perspective. However, superimposing another source of movement on top of the joint movements is not without profound effects on the 2D projection. For example, camera movements in the same global direction as a marker's movement may slow down, cancel or reverse the movement of the corresponding dot with respect to the plane of the projection. Camera movements in an opposing direction may accelerate the dots. Camera movements orthogonal to the movement of a marker may change both the speed and the direction of the corresponding dot. For theories that base the perception of biological motion on some version of 2D to 3D mapping, introducing a changing perspective makes the transformation more difficult. In order to subtract the changes that are due to the camera movement from the changes that are due to the motion of the PLD requires minimally that one is able to differentiate between them. Previous studies that applied projective geometry or trajectory analyses on the dot movements routinely use displays rendered from a fixed point of observation. One study that provided moving points of observation showing PLDs of humans performing gymnastic activities (Jarraya, Amorim & Bardy, 2005) found that, after the PLD disappeared abruptly, anticipation of the upcoming motion pattern was not affected by either panning or tracking movements of the virtual camera. Unlike the first experiment reported in the present work, the authors did not test for possible differences in recognition or response times on a variety of PLD animations. Experiment 1 began simply by testing whether a moving point of observation interferes with recognition or response time in biological motion perception using a relatively wide variety of human motions. There was no motion tracking and fractal analysis; Experiment 1 served as a demonstration that recognition of humans carrying out all sorts of activities in PLDs is fast and accurate even in the absence of most visual "cues" or candidate invariants of projective geometry, and that nesting the point-light trajectories within camera movement induced changes does not slow recognition.

Experiment 1

The first experiment was designed to test the hypothesis that nesting the original PLDs' motion trajectories within changes in velocity resulting from the camera's movements will not impair biological motion perception and, despite the nested structure of the trajectories, recognition will require the same average amount of time. This hypothesis is based on both real-life and cinematographic studies of motion perception. Even though the hypothesis may match a common sense expectation, it poses challenges to both visual processing (organism-generated) and projective geometrical (environment-generated) theories of biological motion perception. The challenge to both types of theories lies in the superposition of another source of motion on the existing trajectories. Within the information processing framework, extracting biological motion from a more complex trajectory pattern requires more computation and resources that is expected to lead to more failures in recognition and longer response time. For the theories that work their way towards recognition from the invariants of projective geometry, the nested trajectories may not necessarily result in failures or longer response times but surely they may demand more undiscovered invariants in order to explain PLD recognition from a continuously changing perspective.

Method

Twenty two students (14 females, 8 males) from the University of Connecticut participated in the study for partial credit in an introductory psychology course. Participants ranged in age from 18 to 22 years. Each of them had normal or corrected to normal vision.

Ten pairs of animations of PLDs of humans carrying out everyday actions (see Table 6 for details) were prepared. Every action was rendered both from a fixed camera position and from a curvilinear moving trajectory orbiting around the target with the camera looking at the center of the action space. Distance from the target was kept the same in both conditions. Animations were randomly selected from a set of 10 with the constraint that each set contained five animations rendered from fixed and five animations rendered from moving point of observation. Animations were presented to the participants seated in front of a laptop computer (with an eye-to-screen distance roughly 60 cm). Observers were instructed to look at the animations and begin an oral description of what they saw on the screen as soon as they formed a stable impression. Regardless of the length of individual animations (mean length 10 s) they were looped so that each presentation lasted for 30 seconds for all 10 animations. Time and fidelity of the responses were determined from audio recordings that were started and stopped in sync with the PLDs in each trial.

Results

Sound recordings were obtained from 220 trials. Only 8 answers were missing or wrong (all of them were first trials). All descriptions correctly identified a human moving on the screen, 201 were correct descriptions, and 11 were unusual but acceptable approximations of the action (e.g., “strange tribal dance” was considered acceptable for imitating an ape). An independent samples Welch t-test indicated no significant difference in response time (in seconds) between fixed perspective (*M* = 7.41, *SD* = 4.87*)* and moving perspective trials (*M* = 7.40, *SD* = 3.59), *t*(210) = 0.02, *p* = .98. There were differences in how fast a particular animation elicited a response, ranging from an average 4.5 *s* for the easiest to an average of 12 *s* for the most difficult, averaged over all participants. Figure 3 shows the interaction of

PLDs and camera conditions. There were also differences in the average response time across participants, ranging from 4 *s* to 14 *s*.

Discussion of Experiment 1

Experiment 1 established that one type of continuously changing perspective (orbiting around the target in the horizontal plane) does not impair or slow down the perception of biological motion. In 96% of the trials, participants recognized the PLDs. Correct responses were given in seven seconds on average. Response time differences were shaped by individual factors (how quickly a given person arrived at the recognition on average) and how difficult a given PLD activity was to perceive correctly. Experiment 1 also tested the viability of recorded oral responses in a recognition task. In most cases, the responses were precise, one or two word utterances, sometimes followed by the addition of more minor details. It became clear however, that it was worth emphasizing in the instructions to start speaking only after arriving at a confident judgment (if only to make the coding of the audio data less laborious).

Experiments With Motion Analysis

The second experiment investigates (1) if the multi-scale fractality of head sway undergoes a change that is related to recognition, and (2) if the time evolution of the multifractal spectra predicts recognition. Experiment 2a repeats Experiment 1 with head tracking and fractal analyses; Experiment 2b introduces disruptions of the PLDs (e.g., vertical camera movements starting from a top view, scrambled point trajectories). The third experiment applies manipulations to both the geometry of PLDs and to the perceiver's experience, such as showing the PLDs multiple times (a), priming the PLDs with congruent or incongruent images of objects (b), and adding contextual changes (c). In Experiments 2 and 3, fractal analysis serves as the method to investigate the underlying multi-scale processes.

General methods and analyses for Experiment 2 and 3.

All PLDs used in the experiments were prepared from motion capture data collected on human subjects performing various actions. The data files were obtained from the Carnegie Mellon University Motion Capture Database. The motion capture data for each PLD was animated and rendered using Blender Open Source 3D graphics application. Each display used 13 markers attached to the major joints (shoulders, elbows, wrists, hips, knees, and ankles) and the head. The 13 markers were represented as small (15 pixel diameter) black discs against a homogeneous white background, with no information about the supporting surface or the horizon. The discs were rendered in orthographic projection to eliminate differential motion parallax and size differences between the discs being closer or further away from the camera. Where not stated otherwise, the PLD animations were projected on a rear-projection screen (DA-LITE, Indiana) using MPlayer movie player application. For each trial, sound recording was started and stopped simultaneously with the visual displays onset and offset times. The recordings were coded manually to determine both the time and content of the oral responses. If not stated otherwise, instructions were to give oral descriptions of the PLDs as soon as participants formed a clear impression. Participants were naïve to the content or the technology of the displays and they did not receive feedback on their performance during the presentation. While seated approximately 2.8 m from the screen, head movements were recorded with a wireless motion tracker (Polhemus, LIBERTY LATUS). The wireless marker was secured to the participant's occiput using an elastic headband. When not stated otherwise, for each trial, recording of head movements with a frequency of 188 Hz started and ended simultaneously with the 30 s long animations, resulting in a time series of 5640 points.

Individual displacements in head sway were calculated by computing the Euclidean distance between each consecutively recorded position. In order to address the time evolution of the structure, the time series of distances were divided into overlapping epochs of 600 data points with a 300 point overlap. Where not stated otherwise this method yielded 17 epochs per trials. The epochs were submitted to Multifractal Detrended Fluctuation Analysis (MF-DFA, Kantelhardt et al., 2002).

MF-DFA

Detrended fluctuation analysis (DFA) gives primary estimates of the fractality due to different long-range temporal correlations of the small and large fluctuations. To assess temporal correlations for different-sized fluctuations, multifractal DFA (MF-DFA) was used. MF-DFA is a generalization of DFA that uses a parameter *q* to selectively emphasize differently sized fluctuations, as

$$F(q,s)=\left[\frac{1}{N_s}\sum_{v=1}^{N_s}\left\{\frac{1}{s}\sum_{i=1}^{s} y\left[(v-1)s+i\right]-y_v(i)^2\right\}^{q/2}\right]^{1/q}, \qquad (1)$$

where y is the integration of the measured time series, y_v is a series of linear fits to y/N_s non-overlapping windows of size s and

$$F(q, s) \sim s^{H(q)}, \qquad (2)$$

where $H(q)$ reflects the temporal correlations. Standard DFA analyses test for monofractality, evaluating Eq. 2 for $q = 2$ only. $H(q)$ reflects larger fluctuations when $q > 2$ and smaller fluctuations when $q < 2$. In contrast to monofractal processes that exhibit all $H(q)$ equal to $H(2)$, multifractal processes exhibit increasing q with non-increasing $H(q)$. Legendre transformation of $H(q)$ yields a range of Holder exponents $\alpha(q)$, whose range specifies the width of a multifractal spectrum. In the subsequent steps of the analyses, the width of the multifractal spectrum was treated as the main predictor for PLD recognition.

Event history analysis

The probability of recognizing and naming the activity in the PLDs was modeled using event history analysis (Allison, 1984; Singer & Willett, 2003). This statistical technique was developed for longitudinal data in which the probability of a discrete event occurring over time is modeled as a function of a set of predictors. In event history analysis both continuous and categorical predictors can be used to model the dependent measure. In this respect, the method is analogous to multiple regression with the exception that the dependent measure in event history analysis is a discrete event that may occur at any point during the observed period. Event history analysis allows both between- subject, time-invariant differences and within-subject, time-varying differences to be used as predictors. Both types of predictors are integrated into a regular, regression-type format. In event history analysis, contribution to model fit is quantified in terms of reductions in the *maximum likelihood* deviance statistic instead of *ordinary least-squares*. The first step of the analysis is to define the *risk set*. The risk set consists of data recorded before a participant experiences the event (recognition). The participant is “at risk” for the occurrence of recognition, regardless of whether it will actually be experienced or not. After experiencing the event (e.g., recognizing and naming the PLD activity correctly), a participant is dropped from the risk set (and thereby, the analysis). As a result, the risk set may change from trial to trial. The changing risk set is usually depicted in risk or *survival plots* where the proportion of participants still in the risk set are plotted as a function of observation times. Only participants still within the risk set contribute to the analyses on a given trial. The second step estimates the hazard, the probability of the event occurring in any given trial. The hazard is estimated as the proportion of participants who experience the event on a particular trial relative to the number of participants at risk. In event history analysis participants who do not experience the event remain in the risk set and, therefore, in the denominator of the hazard. Therefore, event history requires that the predictors simultaneously explain both the timing of the observed events and the failure to observe the event in a single model. As an indicator of emergent changes in multi-scale dynamics, the main predictor in the current work is the width of the multifractal spectrum (the range of values $\alpha(q)$ in the $f(\alpha(q))$ function). For each epoch of each trial, the width values were entered in the event history models as continuous predictors along with the non-interactive predictors such as the particular PLD, subject, geometrical manipulations (camera movements, dot trajectories) and non-geometrical or contextual manipulations (exposure, priming or context).

Experiment 2a

The second experiment repeats the design and analyses of Experiment 1 with the addition of motion tracking of the observer's head movements. It is expected that this replication of Experiment 1 will again yield result in support of hypothesis 1, namely speed and recognition of still and moving points of observation conditions are not expected to differ significantly. The additional dependent measure of head movements allows a test of the hypothesis that recognition is indexed by multi-scale fluctuations. First, movement data from all trials were divided into a pre- and post-recognition parts, and mean multifractal spectrum width was calculated based on MF-DFA estimations. This preliminary analysis tested whether there was a significant change in MF spectrum width after recognition.

In particular, hypothesis 2 is that the long-range temporal structure of head sway will be less self-correlated before recognition, becoming more self-correlated when participants look at PLDs they have already identified. Wider multifractal spectrum curves are expected in the data recorded before recognition as an indication that the system was more open and exhibited less long range dependence (Ihlen & Vereijken, 2010) during exploration. The comparison of pre- and post-recognition informs about the existence of an event-related change in the multi-fractal spectrum. The second phase of analysis tested the possible contribution of this change in the occurrence of the event.

A third hypothesis is inspired by the contribution of multifractal dynamics to the goal-directed reorganization of systems for the self-organization of executive control (Stephen et al., 2012) and searching for puzzle solutions (Stephen & Dixon, 2009). Here, the width of the multi-fractal spectrum is expected to contribute significantly to recognition in successful trials.

Method

Twenty two students (12 females, 10 males) from the University of Connecticut participated in the study for partial credit in an introductory psychology course. Participants ranged in age from 18 to 22 years. Each participant had normal or corrected to normal vision. PLDs were presented on a back lit projector screen. Presentation and instructions were identical to the method used in Experiment 1. In each trial, the time series of 5640 data points was parsed into 600 data point long epochs. Each epoch overlapped with the previous by 300 data points, yielding 17 epochs per trials. Every epoch was submitted to MF-DFA individually. All analyses were conducted according to the general methods section.

Results

Testing Hypothesis 1: Comparison of response times by condition, PLD and participant.

Of the 220 trials, 15 were missed or wrong answers, and 15 were unusual but acceptable descriptions of the actions. Similarly to Experiment 1, there were no significant differences in response time between fixed perspective and moving perspective trials. An independent samples Welch t-test indicated no significant difference in response time (in seconds) between fixed perspective (M = 8.28, SD = 4.98) and moving perspective trials (M = 7.89, SD = 4.28), $t<1$. Individual animations differed in how quickly they elicited a response, ranging from M = 5.8 s for the easiest to M = 11.8 s for the most difficult, averaged over all participants. Figure 4 shows the interactions between PLDs and camera conditions. In addition, average response time differed across participants, ranging from M = 4.5 s to M = 16.4 s.

Testing Hypothesis 2: width of the MF spectrum in pre- and post-recognition data

Analysis of the MF-DFA results showed significantly narrower multi-fractal spectra for the time series collected after recognition. During the process of recognition, mean width (min α – max α) was 1.35; after recognition mean width was 1.17. Welch t test indicated a significant drop $t(2092) = 6.36$, $p < .0001$. There was no significant difference in MF spectrum width between fixed vs. moving point of observations.

Testing Hypothesis 3: event history analysis of recognition

The survival function (shown in Figure 5, for each condition separately) reveals that the proportion of participants still exploring the PLD, looking for an interpretation, decreased systematically as a function of time, similarly for fixed and moving points of observation. The hazard functions for the two conditions (Figure 6) reveal that the probability of recognition increased steadily in epochs 1 through 7 and decreased in epochs 8 – 17, showing that the majority of the events took place in the first half of the presentation.

The effect of MF spectrum width on the probability of PLD recognition was modeled using event history analysis. Besides MF spectrum width, the GLM model included time dependent variables (epoch, trial) and the two camera conditions as categorical predictors. Table 1 summarizes the coefficients for the model. The significant effects of epoch and trial simply indicates that recognition occurred with a higher probability as time progressed. Four PLDs contributed significantly to recognition; those were the PLDs that yielded recognition most frequently. Multifractal spectrum width contributed significantly ($p < .0001$) to event occurrence.

Modeling the effect of the main predictor allows only for estimating the overall strength of its contribution to the event of recognition. In order to trace the effect in slightly more detail, the multifractal width of the data for each epoch was sorted into quartiles. The contribution of each quartile is depicted in a third dimension of a survival plot. Figure 7 illustrates the effect by plotting the survival functions against the four quartiles of width. Regardless of when recognition occurred in time, it was preceded by a relatively large drop in width.

Observation path had no significant effect ($p = .49$). Individual differences in mean response times among participants did not contribute significantly to the model. Time-dependent variables such as epoch and trial contributed significantly as expected from the survival and hazard plots. Most PLDs (except for PLD 5 and 6) contributed significantly showing that there were consistent differences in how difficult a particular PLD was to recognize.

Experiment 2b

In Experiments 1 and 2a the presented PLDs were recognized in nearly every trial (97%). Most recognition took place in first third of the presentation (within about 10 secs). Experiment 2b was designed to challenge recognition by introducing geometrical manipulations in the PLDs to extend the time needed for recognition. This also allows a test of the contribution of MF spectrum width over longer searching times.

PLD recognition can be made more difficult by manipulating the point of observation and the spatial configuration of the dots. Previous work (Sumi, 1984; Pavlova & Sokolov, 2000; Shipley, 2003) has shown that inverted PLDs are nearly impossible to recognize. Pilot work for this dissertation indicated that PLDs rendered according to a top view are difficult to recognize. In Experiment 2b three types of virtual camera positions were used. Each type started from a top view and either remained in that position throughout the entire presentation (top) or gradually approached a 90 degree, horizontal viewing angle on the target in 15 (fast) or in 30 seconds (slow) by moving on a vertical curvilinear path that kept the distance from the target constant. Three presentations out of the ten were scrambled PLDs, where the dots representing the movements of the joints were displaced from the location of the joints by an arbitrary distance (but no longer than one leg length) in a random direction. The three scrambled PLDs were presented randomly in either top, fast or slow types of camera paths.

Hypothesis 1. Unlike virtual camera movements in the horizontal plane, vertical camera movements starting from a top view are expected to affect the observer's response times. The top view is an observation point that is not commonly available in everyday situations. It is hypothesized that the faster the PLD departs from a top view the easier it will be recognized.

Hypothesis 2. Even though the dots removed from the joint locations still trace the movement of the individual joints, the resultant PLD fails to carry the global transformations of a human figure, making it extremely difficult to recognize. Scrambled PLDs were expected to yield no recognition regardless of the path of observation.

Hypothesis 3. Despite the applied geometrical difficulties, recognition should still be tied to the interaction of many participating scales. The width of the multi-fractal spectrum is expected to remain the main contributor to recognition in all successful trials.

Both response times and the transition to more long-range dependence (less intermittency) in head sway is expected to correlate with the departure from the “difficult” viewing angles.

Methods

Seventeen students (12 females, 5 males) from the University of Connecticut participated in the study for partial credit in an introductory psychology course. Participants ranged in age from 18 to 22 years. Each participant had normal or corrected to normal vision. Presentation and data analyses were conducted according to the general methods section. For every participant, ten PLD animations were presented. The three types observation paths (top, slow and fast) and the two types of PLD configurations (normal and scrambled) were presented in a random order. Throughout the experiment both manipulations were counterbalanced; all conditions were equally represented. Participants were asked to maintain a comfortable seated posture and identify the displays orally as soon as they arrived at a confident recognition.

Results

In 55 trials out of the 170, participants were able to give a correct description. Scrambled PLDs included, participants achieved recognition in 32% of the trials. Excluding the nearly impossible scrambled trials, performance was successful in 43% of the trials.

Testing Hypothesis 1: Digression from top view enhances the probability of recognition

Response times were not significantly longer for top view conditions (*M* = 21 s) than for either slow (*M* = 21.3 s) or fast descending views (*M* = 20.6 s). According to a within subject analysis of variance, the main effect of observation path on response times was non-significant, *F*(2, 655) = 1.76, *p* = 0.34. It should be noted that response times can only be reported in successful trials; consequently this analysis excludes cases when there was no recognition. A better way of measuring performance, therefore, is to look at the percentage of correct responses. Top view PLDs were recognized on 22% of trials, slowly descending PLDs 45% and fast descending PLDs 45% of the times in the non-scrambled presentations. A within subject analysis of variance showed a significant main effect for observation path on the percentage of correct responses, *F*(2, 166) = 3.6, *p* < .023. Multiple comparisons showed that both slow and fast descending conditions yielded significantly higher portion of correct responses than the top view condition (*p* < .0001) but there was no significant difference between slow and fast conditions (*p* = .1, see figure 8).

Testing Hypothesis 2: Scrambled PLDs inhibit recognition

As expected, relocating the markers impaired PLD recognition severely.

Testing Hypothesis 3: Width of the multi-fractal spectrum predicts the event of recognition

Event history analyses confirmed some of the previous findings showing similar survival curves for the camera movement types and no event in the scrambled condition (Figure 9 a and b). GLM modeling included MF spectrum width, time dependent factors (epoch, trial), PLDs, observation path types (top, slow, fast) and PLD configurations (normal, scrambled) as categorical predictors). Table 2 summarizes the coefficients for the model. The significance of epoch indicates that recognition became more likely as time progressed. Neither trials nor PLDs were significant predictors. Observation path types were moderate contributors for the occurrence of the event. The fact that observation types did not contribute to recognition was confirmed by a second model that included only successful trials. As expected, the effect of the three observation types disappeared from the second model. In both models, MF spectrum width was the strongest contributor ($p < .0001$). Similarly to Experiment 2a, a surface plot (see Figure 10) of the survival functions by MF width quartiles showed the largest drops in the fourth quartile, albeit for longer curves than the ones obtained in Experiment 2a.

Discussion of Experiment 2a and 2b

Experiment 2a provided further support for the hypothesis that parallel projection of human biological motion according to fixed and horizontally moving points of observation are recognized and reported within the same mean response times. A coarse pre- and post recognition multifractal analysis of head sway showed a significant difference in the long-range dependencies of different size movements, indicating less intermittency in the time-series collected after recognition. Event history analysis on subsequent epochs of head movements confirmed the hypothesis that MF spectrum width plays contributes significantly to recognition. This latter finding indicates that visual recognition is contingent upon a specific evolution of interaction among fluctuations of different sizes. These fluctuations are considered as indexes of multiplicative multifractal cascade dynamics in between the participating levels of organization both in the environment and in the organism.

Experiment 2b introduced two types of geometrical manipulations of the PLDs; vertical camera movements starting from a top view on the target, and scrambled point trajectories. Both manipulations affected the observers' performance in identifying PLD activities. Scrambled PLDs were nearly impossible to recognize, regardless of camera movements. For non-scrambled PLDs the number of correct responses was lowest for the fixed top view presentation and highest for quickly descending camera condition. Interestingly, this manipulation had no effect on average response times; however, the average response times were higher (20 s) than the response times in Experiment 1 and 2a (8 s). Regardless of how strong the effect of the projective manipulations was in Experiment 2b, MF spectrum width remained the strongest time varying predictor of the event of recognition.

In Experiment 3, various non-geometrical manipulations were tested in conjunction with the geometrical manipulations of Experiment 2. More specifically, in Experiment 3, besides shaping the geometry of the projections, the perceiver's experience will be manipulated by factors such as: (a) showing the PLDs multiple times, (b) priming the PLD recognition with congruent or incongruent images of objects, and (c) applying contextual changes in the order of the presentations.

Experiment 3a

In Experiment 2, recognition of human biological motion was tested in conditions that differed only in terms of the use of projective geometry. The directions and velocities of a target-centered moving camera and the arbitrary displacement of the markers created an extra source of change in the dot trajectories over and above the changes due to the movement of the joints. Some of these changes impaired the observers' perception, others seemed to have no effect on their performance. Despite the degree of imposed difficulties, however, whenever correct identification occurred it was best predicted by the width of the MF spectrum of head sway. According to the interactive approach to perception, the nested fluctuation of energy patterns in the optic flow is but one source of influence that feeds into the self-organization of recognition. Another equally important source is the organism being busy with scanning the optic flow. Perceptual systems are never at rest; tremor, jitter, eye movements, head turns, postural adjustments, involuntary and voluntary movements occur at all participating scales. Moreover, perceptual systems are nested within the rest of the biological "machinery" working at dissipating, storing and utilizing energy in all ongoing life sustaining processes. Some of these processes feed cognition, memories and other ways that help the organism to situate itself in the environment. Presumably, identification of a particular human activity through its PLD presentation is more difficult for observers who never experienced a certain activity by either seeing others engaged in it or carrying out the activity themselves. In these cases, the self-organization of recognition may miss something but that miss is not due to the geometry of the projection. Experiences, expertise, familiarity, education of attention, context and situation play key roles in perception.

There is indication that experience and feedback induce changes across perceptual systems (Michaels, Arzamarski, Isenhower & Jacobs, 2008 ; Withagen & Michaels, 2005). In perceptual learning, exploration and feedback are expected to drive observers across an information manifold towards loci that help them differentiate more precisely in the available energy patterns with respect to their goals (Jacobs & Michaels, 2007; Michaels et al., 2008). These factors are themselves tied to ongoing processes within the perceiver. The idea here is that factors usually attributed to cognitive processing are supported by and participate in the same multiplicative cascade dynamics as, say, the movements of the eye or the head. Studied and conceptualized in seemingly distant spatial and temporal domains from one another, there is not a single process in and around the body that could run physically isolated from the scheme that Gottlieb (1998, 2000) envisioned. The general idea in Experiment 3a is to apply geometrical manipulations in conjunction with multiple exposures and to test if the contribution of change in the fractal scaling holds up when both geometrical and non-geometrical factors shape the probability of perception.

Hypothesis 1. Digression from top view enhances the probability of recognition

Experiment 3a used PLDs from the same set of animations presented in Experiment 2b. It was hypothesized that even if multiple exposure enhance the probability of recognition, the overall effect of observation path should remain a contributor. PLDs providing a descending point of observation are expected to lead to the event of recognition more often than the ones fixed in the top view.

Hypothesis 2. Previous exposure to the same PLD enhances the probability of recognition even in the absence of feedback

Previous exposure is expected to enhance performance in two ways. First, the second encounter may reveal more about the same activity in terms of providing more time to explore and providing a different path of observation. Second, if the previous encounter yields a correct identification, the physical similarities of the fluctuating pattern may enable participants to identify the action correctly form a point or path of observation that would have not led to recognition as a first encounter. Even though this second effect might not arise because the presentations become progressively more difficult (e.g. fast –> slow -> top), the overall effect of multiple exposure is expected to enhance performance.

Hypothesis 3. Width of the multi-fractal spectrum predicts the event of recognition Regardless of the origin or the time scale of the multifractal fluctuations that support it, recognition is supposed to be tied to the interaction of many participating scales. The width of the multi-fractal spectrum was expected to remain the main contributor to the event of recognition. Transitions to more long range dependence (less intermittency) in head sway is expected to positively correlate with the digression from top view and the number of exposures to the same PLD activities.

Method

Eighteen students (10 females, 8 males) from the University of Connecticut participated in the study for partial credit in an introductory psychology course. Participants ranged in age from 18 to 22 years. Each participant had normal or corrected to normal vision. Presentation and data analyses were conducted according to the general methods section. For every participant, fifteen animations were presented in random order from a set of five PLDs. There were three types observation paths: top view, descending to the side view slowly (in 30 seconds), descending to side view fast (in 15 seconds). Each participant encountered every PLD three times, albeit in three different observation paths. The order of presentation was 1-1-1-1-2-2-2-1-2-3-3-2-3-3-3, where the numbers represent the number of times the observers encountered a PLD presentation. The PLDs were randomly assigned to the fixed order of presentation. Participants were asked to maintain a comfortable seated posture and identify the displays orally as soon as they arrived at a stable, confident recognition. No feedback was provided in between the presentation of the same PLD.

Results

Participants reported correct identifications of the perceived activities 43% of the trials. Similarly to Experiment 2b, reporting response times would have limited the analyses to successful trials only, therefore proportions of correct responses are reported.

Testing Hypothesis 1: Digression from top view enhances the probability of recognition

Top view PLDs were recognized on 12% of the trials, slowly descending PLDs 35% and fast descending PLDs 52%. Figure 11 summarizes these findings. According to a within-subject analysis of variance, the main effect of observation path on the percentage of correct responses was significant, $F(2, 267) = 19.72$, $p < .0001$. Planned comparisons of the means revealed highly significant ($p < .0001$) differences between all three viewing conditions.

Testing Hypothesis 2: Previous exposure enhances the probability of recognition

Main effect of exposure was significant, $F(2, 267) = 8.88$, $p < .001$. First encounters yielded correct responses in 18% of the trials, while second and third encounters yielded 44% and 34% of the trials (see Figure 12) respectively. Multiple comparisons revealed significant differences between the first and the other two levels of encounters ($p < .001$). Surprisingly, second encounters elicited a higher percentage of correct response than third encounters, however, the difference was not significant.

Testing Hypothesis 3: Width of the multi-fractal spectrum predicts the event of recognition

Event history analyses confirmed the findings reported in testing hypotheses 1 and 2, showing the earliest decline for the “fast”, moderate decline for the “slow” and latest decline for the survival curve of the “top view” condition (Figure 13a). The effect of exposure was also confirmed by the longest tail for the first encounter and progressively shorter tails for second and third encounters (Figure 13b). GLM modeling included MF spectrum width, time dependent factors (epoch, trial), PLDs, observation path types (top, slow, fast) and exposure (first, second, third) as categorical predictors. Table 3 summarizes the coefficients for the model. The contribution of epoch indicated time dependence. The two animations that were most difficult to recognize five predicted the probability of recognition significantly and negatively. Both observation path types and exposure had significant effects. MF spectrum width once again predicted the occurrence of the event ($p < .001$). Path types, exposure and MF spectrum width showed no significant interaction, and including interaction terms did not improve the model fits (Chi square test of the residual variances was not significant). A surface plot (see Figure 14) of the survival functions by MS width quartiles showed the largest drops between the first and second quartile.

Discussion of Experiment 3a

In Experiment 3a five PLDs were presented according to three paths of observation (top, slow, fast). Every PLD action was presented three times during a session, albeit, according to different paths of observation. The results indicate that both geometrical and non-geometrical manipulations affected recognition. Both observation paths digressing from the top view and the number of encounters enhanced the probability of recognition. Despite the significant effects originating from radically different sources, the width of the MF spectrum of head sway remained a strong predictor of the event of recognition. These findings suggest that visual recognition is contingent upon the interaction of fluctuating scales originating both in the patterns of optic flow and within the perceptual system of the observer.

Experiment 3b

Recognizing an event under real life circumstances entails responding to the contingencies that make up the event. In visual recognition, those contingencies include forms and transformations of various kinds reoccurring together under similar conditions. For example, recognizing an activity of a person may be helped by recognizing her outfit or the tools she uses besides identifying motion patterns. In PLD animations these kinds of contingencies are absent but, perhaps they can be added by “priming” the recognition with images of clothing or tools that are normally used in the presented activity. With this kind of manipulation, another level of optical information was supplied in conjunction with the geometrical manipulations in the PLDs.

In Experiment 3b, pairs of visual stimuli were created from a still image of an object and a PLD animation. The images were either strongly related or unrelated to the PLD activity in each pair. For example, if the PLD was a human figure sweeping the floor, either a broom was shown (congruent case) or a pair of scissors (incongruent case).

Hypothesis 1: Digression from top view enhances the probability of recognition

Experiment 3b also used PLDs from the same set of animations presented in Experiment 2b. It was hypothesized that even if the congruency of priming images enhances the probability of recognition, the overall effect of observation path will remain a contributor. PLDs providing a descending point of observation are expected to lead to recognition more often than the ones fixed in top view.

Hypothesis 2: Congruent images will enhance the probability of recognition

Photographs and images in general shape the fluctuation of light scattering as a result of the light interacting with the surface of the displays. The technique or mechanism used for creating the images results in fluctuating patterns that are more or less similar to the energy flow of the scene the photograph was taken from (Gibson, 1971). According to Gibson's

(1979) theory of affordances, objects are perceived in terms of action possibilities. In the PLDs, however, the tools themselves were not indicated by the points. Whether or not the fit between action and object influenced event recognition was addressed through a priming paradigm. In the congruent condition, images of tools were chosen to match the action presented in the upcoming PLD, in order to provide extra information on action possibilities. Therefore, congruent images were expected to enhance the probability of recognition.

Hypothesis 3: Width of the multi-fractal spectrum predicts the event of recognition

Multifractal fluctuations in the optic flow originate in movement of many different scales from wavelength size light-surface interactions to large scale movements of objects, and from the small scale (e.g., ocular tremor) vibration to the large scale (locomotion) movements of the perceiver. Every participating scale “enters” into the self organization of recognition. Therefore, the width of the multi-fractal spectrum, indicating that recognition is a result of fluctuations originated in the presented images, the PLDs and the movements of the observers, is expected to remain the main contributor to the event of recognition in all successful trials. Transitions toward more long range dependence in head sway is expected to positively correlate with the digression from top view and the congruency of the priming images, therefore, with the event of recognition.

Method

Twenty students (12 females, 8 males) from the University of Connecticut participated in the study for partial credit in an introductory psychology course. Participants ranged in age from 18 to 22 years. Each participant had normal or corrected to normal vision. Presentation and data analyses were conducted according to the general methods section. Each participant observed ten PLD animations. There were two types observation paths: top view and descending to side view fast (in 15 seconds). Before the onset of each PLD a still image of an object was presented for five seconds. The object was either strongly related or unrelated to the activity. Participants were asked to maintain a comfortable seated posture, observe the still images preceding the PLDs and identify the displays orally as soon as they arrived at a stable, confident recognition. To assure attention to the primes, they were told to scrutinize the images as if they had to draw a sketch of them later.

Results

Participants correctly identified the perceived activities on 57% of the trials. As a measure of performance, percentages of correct responses are reported.

Testing Hypothesis 1: Digression from top view enhances the probability of recognition

Top view PLDs were recognized 45%, fast descending PLDs 69% of the presentation times. Figure 15a summarizes these findings. According to a within subject analysis of variance, the effect of observation path on the percentage of correct responses was significant, $F(1, 198) = 12.36$, $p < .001$.

Testing Hypothesis 2: Congruent images will enhance the probability of recognition

Incongruent image-PLD pairs yielded correct responses in 49%, while congruent image-PLD pairs resulted in correct responses in 65% of the trials (see Figure 15b). The effect of congruency was significant, $F(1, 198) = 5.5$, $p = .02$.

Testing Hypothesis 3: Width of the multi-fractal spectrum predicts the event of recognition

Event history analyses confirmed the differences reported above, showing the earlier decline for the “fast”, and later decline for “top view” condition (Figure 16a). The effect of congruency was also confirmed by the longer “survivals” for the incongruent pairs of presentations (Figure 16b). GLM modeling included MF spectrum width, time dependent factors (epoch, trial), PLDs, observation path types (top, fast) and congruency as categorical predictors. Table 4 summarizes the coefficients for the model. The significant contribution of epoch and trial indicated time dependence. PLDs that were recognized the most frequently contributed the significantly to the model. Observation path types made modestly significant contribution; congruency made significant contribution. MF spectrum width once again predicted the occurrence of the event ($p < .0001$). A surface plot (see Figure 17) of the survival functions by MF width quartiles showed the largest drops between the second and third quartiles.

Discussion of Experiment 3b

In Experiment 3b ten PLDs were presented according to two paths of observation (top, fast). Every PLD animation was preceded by a brief presentation of a photograph of an object either congruent or incongruent with the human activity. The results indicate that both geometrical and priming manipulations affected recognition. Both observation paths digressing from the top view and the exposure of congruent priming images enhanced the probability of recognition. Despite the significant effects originating in radically different sources, the width of the MF spectrum of head sway remained a strong predictor of the event of recognition. The positive effect of images suggests that action possibilities perceived in the images of tools facilitated recognition of the action in the congruent condition. According to ecological accounts of perceptual learning and development (Gibson & Pick, 2000) being able to perceive action possibilities requires previous encounters with these activities in information rich, natural situations. The positive effect of priming opens the possibility that visual recognition is contingent upon the interaction of fluctuating scales originating both in the patterns of optic flow and within the perceptual system of the observer even outside of the timescale of the presentation.

Experiment 3c

Experiment 3c tests the overall hypothesis that the perception of similar events, or events that belong to the same behavioral context do not require the same degree of reorganization in the interaction dynamics, therefore the recognition of such events may show higher probability than events that belong to a different context. The guiding hypothesis of Experiment 3c is that the behavioral context of the presented PLDs plays a key role in their recognition.

Hypothesis 1: The probability of recognition depends on the point of observation

Experiment 3c also used PLDs from Experiment 2a and Experiment 2b. There were two types of camera movements, one was orbiting around the target in the horizontal plane and the other was descending fast from the top view. In order to collect a sufficient number of animations that fit the two types of activities, PLDs with horizontal camera movement from Experiment 1 and 2 were included. It was hypothesized that even if the contextual effect in the order of presentation enhances the probability of recognition the overall effect of observation path remains a contributor. PLDs rendered according to a side view were expected to lead to the event of recognition more often than the ones descending from top view.

Hypothesis 2: Less frequent contextual change enhances the probability of recognition

Subsequent presentations of PLDs that belonged to the same behavioral context will be recognized more often than the ones following a PLD of a different behavioral context.

Hypothesis 3: Width of the multi-fractal spectrum predicts the event of recognition

If the recognition of events that belong to the same behavioral context requires less reorganization in the interaction dynamics, they may occur with a higher frequency. Once again, changes in the width of the MF spectrum are expected to contribute to recognition.

Method

Eighteen students (14 females, 4 males) were recruited from the University of Connecticut for partial credit in an introductory psychology course. Participants ranged in age from 18 to 22 years. Each participant had normal or corrected to normal vision. Each participant observed ten PLD animations of humans engaging in either a "sporting" or a "working" activity. One PLD showed a human sitting down and standing up repeatedly, this one was a neutral activity. Three PLDs were rendered according to a side view, seven PLDs were rendered according to a quickly descending path of observation from the top view to the side. The order of the animations was manipulated to change the context from sporting to working with few or many switches ranging from three to eight times. The fewer times the context was changed, the longer the participants had to stay within the same category of activity.

Results

Participants reported correct identifications of the perceived activities on 57% of the trials. As a measure of performance, percentages of correct responses are reported.

Testing Hypothesis 1: Path of observation has an effect on the probability of recognition Side view PLDs were recognized 94%, fast descending PLDs were recognized 43% of the presentation times. Figure 18a summarizes these findings. According to a within subject analysis of variance, the effect of observation path on the percentage of correct responses was significant, $F(1, 178) = 53.15$, $p < .0001$.

Testing Hypothesis 2: Less frequent contextual change enhances *the probability of recognition*

PLDs presented within the context of the previous PLD yielded correct responses in 45%, while changing the context with respect to the previous presentation resulted in correct responses in 39% of the trials (see Figure 18b). The effect of contextual change was significant, $F(1, 178) = 3.52$, $p = .03$.

Testing Hypothesis 3: Width of the multi-fractal spectrum predicts the event of recognition

Event history analyses confirmed the differences reported above, showing a much earlier decline for the side view condition than the decline for the “fast” condition (Figure 19a). The effect of changing context was also confirmed by the longer “survivals” for PLDs belonging to a different context than the previous PLD (Figure 19b). GLM modeling included MF spectrum width, time dependent factors (epoch, trial), PLDs, observation path types (side view, fast) and contextual change as categorical predictors. Table 5 summarizes the coefficients for the model. The significant effect of epoch indicated time dependence. The contribution of PLDs were highest for PLDs that were recognized most frequently. MF spectrum width once again predicted the occurrence of the event ($p < .0001$). A surface plot (see Figure 20) of the survival functions by MS width quartiles showed the largest drops between the second and third quartiles.

Discussion of Experiments 3c

Experiment 3c investigated whether changing the behavioral context of subsequent PLD presentation has a negative effect on the efficacy of recognition. In this experiment the modest aim was to demonstrate that obvious contextual changes that are defined by the experimenter have an effect on the probability of recognition. The choice of behavioral contexts (sport, work) was pretty much dictated by the set of animations that were prepared for all the experiments. However, the result indicated that even these contextual categories that were assembled ad hoc from an existing set of PLDs had a significant effect on perception. One way of interpreting the results is that when the context changes, the dynamics of the interaction has to undergo a re-organization that takes a certain amount of time and it temporarily sets back perceptual performance.

General Discussion

The present dissertation sought ways of addressing the “work” that is being done in order to perceive the visual world. In a conceptual framework that is radically different from traditional accounts, it is argued that this work should be attributed to neither the organism nor to the environment alone. The approach pursued here seeks explanations in the interactions of environmental and organismic processes, processes and events that are spread out through a vast and barely known terrain of multiplicative cascade dynamics. The conceptual foundation of this work is a proposed union between the Ecological approach of the late James Gibson (1979) and the theory of Probabilistic Epigenesis of the late Gilbert Gottlieb (1998, 2000). According to Gibson, there is no organism without an environment, and there is no environment without an organism. They imply one another and the co-implication is a result of billions of years of evolution. Perception is but one among various ways living systems sustain themselves by exchanging energy and matter with their environment. Perception is rooted in the exchange carried out on a massively distributed, nested network of nonlinear processes ranging from molecular scales to the ecosystem level.

In Gottlieb's model (Figure 1), every process is nested within larger and/or slower processes and made up of smaller and/or faster processes. Every relationship is bidirectional, processes are both forcing and reacting to other processes both horizontally and vertically in the model's plane. Mechanical, electrical, chemical, and other forms of energy are stored and dissipated through the network and the transports are indexed by fluctuations on all scales (Van Orden, Holden & Turvey, 2003, 2005; Stephen & Dixon, 2009).

From environmental gradients, such as electromagnetic radiation, energy flows into perceptual systems, building up and interacting with gradients that are sustained there. Depending on frequency and resolution, a measurement taken somewhere in the perceptual system captures more or fewer of the interactions. Just like the patterned energy that feeds into the perceptual system, the measurement has a nested multi-scale structure best characterized as multifractal. Multifractal structure is a result of the numerous scales of simultaneous activity that make up the system. Analyses of these measurements have been used profitably to predict changes in perceptual and cognitive behavior (Dixon, Holden, Mirman & Stephen, 2012). This dissertation applies this method to the problem of visual perception.

The task

The choice of visual task was biological motion perception. There are a number of factors that make point-light displays of humans carrying out common activities a suitable vehicle for present purposes. First, it is a display that is based on a measurement taken on the joints of a human actor. Even though much of the richness of human movements is bleached out for simplicity and ease of use for computer animation, it is still a pattern that originated on a biological system. Second, ever since they were first created, PLDs have resisted traditional theories and explanations of visual perception. Theories that wish to use pictorial cues are left with almost none. Theories that wish to use integration of still images are left with incomprehensible frames to begin with. The proposed invariants of projective geometry can be weakened or eliminated without affecting the perception of PLDs. The third advantage is that PLDs are relatively easy to manipulate in order to target specific experimental questions and create the types of variations that were used throughout the experiments. Finally, PLDs can slow perception sufficiently so as to allow for application of the time series analyses relevant for assessing multi-scale dynamics.

Experiment 1

The first experiment established that PLD recognition remains robust and accurate even when the point-light trajectories due to the movement of the joints are embedded within the trajectories due to camera movements. The purpose of this demonstration was to raise awareness that after four decades of research, PLDs are still challenging to perceptual theories and they can be made more challenging for the theories without making perception any more challenging to observers. Therefore, Experiment 1 also served as a suitable entry point for the alternative framework considered in the present work.

Experiment 2

Experiment 2 introduced the methods and analyses used throughout the dissertation. Epochs of head sway data were submitted to multifractal analysis in order to estimate the width of the MF spectrum for each epoch. This measure, together with the levels of viewing conditions, were submitted to event history analysis to model the contribution of the predictors to recognition. Event history analysis revealed a highly significant contribution of MF spectral width. This finding is consistent with the hypothesis that the evolution of the MF spectrum width indicate changes in the dynamics of exploration. Experiment 2 found that exploration, - a search for a solution or interpretation - exhibits more intermittency, openness or sensitivity to influences, as evidenced by less long-range dependency. But as soon as the system gets progressively closer to recognition, long-range dependency becomes stronger and the strength is indicated by the narrowing of the spectrum. One might imagine tracing the path of an animal, searching for resources and finally finding them. The path contains many turns, shorter and longer distances covered in many directions during the search, but eventually the path becomes more predictable as the animal makes a transition from exploration to homing on its target. The dynamics of search behavior has been shown to indicate multiplicative interactions, not only for foraging animals (Shlesinger, Zaslavsky & Klafter, 1993), but for eye-movements (Stephen & Mirman, 2010) and even memory search (Rhodes & Turvey, 2007). The present findings strongly suggest that success in search may be predicted from the changes in multi-scale dynamics.

Experiment 3

The third experiment kept the geometrical manipulations of the PLDs and introduced non-geometrical factors such as multiple exposure, priming with images, and switches of behavioral context. All three manipulations were meant to evoke changes that are usually conceptualized as higher-level data processing that is isolated, abstract and somewhat removed from the dynamics of ongoing life processes. For example, multiple exposure was meant to invoke the concept of memory. If an animal encounters event E at some point in time, then it encounters some events different from E, and at some point in time E is experienced again and the animal's response time or performance is better than it was at the first encounter, the concepts of learning or memory are typically called for. However, experiencing the same event from different perspectives is the norm for all growing and locomoting life forms. Sometimes the continuous (static or changing) perspective on E may get interrupted by obstacles or detours resulting in shorter or longer periods of absence of E. There is no principled way of defining the demarcation line between perception and memory and perhaps there is no need to do so. From an interaction-dominant point of view, fluctuations are quasi-periodic changes in intensity and frequency with no built-in limits for either. The results of Experiment 3a indicated that a second exposure to the same pattern enhanced the probability of recognition.

In Experiments 3b and 3c, the effect of visual and behavioral context was tested. Relating events to their contexts is a central theme in the interactive approach. In ecological psychology, behavior arises from and is specific to the context in which it is nested (Gibson, 1979). Context is a concept notoriously difficult to define. It usually gets described as a set of circumstances that are relevant to some actuality in behavior, but this description is vague, because relevance is also contingent upon the context. The idea of a change in the context, therefore, is no less ambiguous but perhaps, the concept still allows for more or less obvious cases for contextual changes. For example, in Experiment 3b, congruent images helped to set the stage for a higher number of correct responses than incongruent images. In Experiment 3c, changing the context in a series of presentations resulted in fewer correct responses. Behavioral-, or environmental contexts, however, may not always line up very well with our expectations. What “matters” --what is relevant to a given situation-- is also influenced by contextual factors and it may change in non-obvious ways.

Conclusions

This dissertation proposed multiplicative multifractal cascade dynamics as a general framework for visual recognition. The act of exploration of a visual scene and the recognition of action or a change in action has been considered as part of an ongoing multi-scale interaction between the perceiver and the perceived. As such, this approach contributes to a prominent topic in cognitive science—what does it mean for perception and cognition to be embodied and embedded. Existing theories of vision tend to place emphasis on either processes within the organism (e.g., inferential processes in the visual cortex) or invariants on a specific scale in the visual environment (e.g., time-to-contact, cross-ratio). However, the choice of geometry is motivated by theoretical assumptions about the nature of the visual problem at hand and by the mathematical tools used to address them. As a result, there are many different types of geometries proposed for particular visual problems. Euclidean, algebraic, differential or projective geometries are routinely applied to address specific problems in vision such as figure-ground differentiation, object or face recognition, contour-integration, or motion perception. The diversity of solutions hinders theoretical integration and generalization across species with comparable visually guided behavior but different perceptual and nervous systems. In contrast, the framework proposed here allows the geometry to be dictated by the natural structure and evolution of the participating processes. It offers a more task-general and less anatomically driven path to the problem of recognition that may generalize across species and across senses.

The proposed framework is compatible with and motivated by modern physical and biological descriptions of self-organizing natural systems. This compatibility permits and encourages seeking further theoretical connections with complex systems behavior. Multiplicative cascade dynamics are closely related to stochastic chaotic behavior of turbulent systems (Schertzer & Lovejoy, 2004). Turbulent systems exhibit transitions to laminar flow when they lose kinetic energy due to molecular viscosity. Testing whether a transition from exploration to recognition reported in the present work is related to a transition from turbulent flow to laminar flow may provide further insights into the relation of cognitive behavior and cascade dynamics.

The Gottliebian, multi-scale framework is consistent with Gibson's ecological approach to visual perception in terms of emphasizing organism-environment mutuality; it considers the senses as perceptual systems and perception as inseparable from action. It offers physical underpinnings to direct perception by rejecting the idea of distinguished scales or loci of processing visual information, even in perceptual relations that were traditionally labeled as indirect. Although, Gibson's theory of affordances was not directly addressed by the present work, the method present allows future investigations in that domain. In particular, the recognition of action possibilities with respect to a given or freely chosen goal may also be indexed by changes of multifractal spectrum width in eye-movements, head sway or similar motion capture data.

The design and the analytical tools (MF-DFA and event history analysis) revealed surprisingly reliable and strong contributions of MF spectrum changes to recognition in a variety of circumstances. The robustness of the findings invite further investigations in various types of visual and non-visual recognition and search behavior in general.

The emphasis on multi-scale interactions in visual perception provides an alternative framework to visual perception and may lead to novel technologies in creating visual aids, artificial bionic or mechanical visual devices. Moreover, these experiments opened up new territories to explore: relating visual experience and behavior to measures in the interaction dynamics within an undivided organism-environment system.

References

Ahlstrőm, V., Blake, R., & Ahlstrőm, U. (1997). Perception of biological motion. *Perception*.

Aks, D. J., Zelinsky, G. J., & Sprott, J. C. (2002). Memory across eye-movements: 1/f dynamic in visual search. *Nonlinear dynamics, psychology, and life sciences*, *6*(1), 1-25.

Allison, P. D. (Ed.). (1984). *Event history analysis: Regression for longitudinal event data* (No. 46). Sage.

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. Perception and Psychophysics, 23, 145–152.

Beardsworth, T., & Buckner, T. (1981). The ability to recognize oneself from a video recording of one's movements without seeing one's body. *Bulletin of the Psychonomic Society*.

Beintema, J.A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences, 99,* 5661-5663

Bertenthal, B.I., Proffitt, D.R., and Cutting, J.E. (1984). Infant sensitivity to figural coherence in biomechanical motions. *J. Exp. Child Psychol. 37,* 213–230.

Bertenthal, B.I., and Pinto, J. (1994) Global processing of biological motions. *Psychol. Sci. 5,* 221–225.

Bingham, G.P. (1993). Scaling judgements of lifting weights: Lifter size and the role of the standard. *Ecological Psychology, 5,* 31–64.

Blake, R. (1993). Cats perceive biological motion. *Psychological Science, 4,* 54-57.

Bronstein, A.M. & Buckwell, D. (1997) Automatic control of postural sway by visual motion parallax. *Experimental Brain Research, 113,* 243–248.

Cutting, J. E. (1986). *Perception with an eye for motion*. Cambridge, MA: Mit Press.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society, 9*, 353-356.

Ding, H., Wang, Z., Nguyen, F., Boppart, S. A., & Popescu, G. (2008). Fourier Transform Light Scattering Of Inhomogeneous And Dynamic Structures. *Physical Review Letters*, *101*(23), 8102.

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature, 170(4314)*, 36-37.

Dittrich, W. H., Troscianko, T., Lea, S., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*, 727-738.

Dixon, J. A., Holden, J. G., Mirman, D., & Stephen, D. G. (2012). Multifractal dynamics in the emergence of cognitive structure. Topics in Cognitive Science, 4, 51-62.

Duarte, M., & Zatsiorsky, V. M. (2000). On the fractal properties of natural human standing. *Neuroscience Letters*, *283*(3), 173-176.

Ehrenfield, T., Guerraz, M., Thilo, K. V., Yardley, L., & Gresty, M. A. (2003). Posture and mental task performance when viewing a moving visual field. *Cognitive Brain Research, 17*, 140–153.

Field, D. J. (1987). Relations between statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4,* 2379–2394.

Gibson, E. J., & Pick, A. D. (2003). *An ecological approach to perceptual learning and development.* New York: Oxford University Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1971). The information available in pictures. *Leonardo, 4,* 27-35.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Giese M. A. & Poggio, T. (2003). Neural mechanisms for the recognition of biological motion. *Nature Reviews Neuroscience, 4*(3), 179–192

Gottlieb, G. (1998). Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesis. *Psychological Review, 105*(4), 792-802.

Gottlieb, G. (2000). Nature and nurture theories. In A. E. Kazdin (Ed.) , *Encyclopedia of psychology, Vol. 5* (pp. 402-404). Washington, DC New York, N

Hahn, D. W. (2006). Light scattering theory. *Department of Mechanical and Aerospace Engineering, University of Florida*.

Hausdorff, J. M., Purdon, P. L., Peng, C.-K., Ladin, Z., Wei, J. Y., & Goldberger, A. L. (1996). Fractal dynamics of human gait: Stability of long-range correlation in stride interval fluctuations. *Journal of Applied Physiology, 80*, 1448-1457.

Hermann, S. (2005). Exploring sitting posture and discomfort using nonlinear analysis methods. *IEEE Transactions on Information Technology in Biomedicine, 9*, 392-401.

Hirai, M., Kaneoke, Y., Nakata, H., & Ryusukekakigi. (2008). Neural Responses Related To Point-light Walker Perception: A Magnetoencephalographic Study. *Clinical Neurophysiology, 119*(12), 2775-2784.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. *NASA STI/Recon Technical Report A, 87*, 33516.

Hodges, B. & Lindheim, O. (2006). Carrying babies and groceries: The effect of moral and social weight on caring, *Ecological Psychology, 16,* 93-111.

Hoffman, D.D. & Flinchbaugh, B.E. (1982). The interpretation of biological motion. *Biological Cybernetics, 42*, 195-204

Ihlen, E. A. F., & Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond 1/*f* fluctuation. *Journal of Experimental Psychology: General*, *139*, 436-463.

Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Stanley, H. E., Struzik, Z. R. (2001). From 1/f noise to multifractal cascades in heartbeat dynamics. *Chaos*, *11*, 641-652.

Jacobs, D. M., & Michaels, C. F. (2007). Direct learning. *Ecological Psychology, 19,* 321-349.

James, W. (1890). *The principles of psychology*. H. Holt and company, New York.

Jarraya, M., Amorim, M. A., & Bardy, B. G. (2005). Optical flow and viewpoint change modulate the perception and memorization of complex motion. *Perception & psychophysics*, *67*(6), 951-961.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception &. Psychophysics, 14,* 201–211.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A, 316,* 87-114.

Keesey, U. T. (1960). Effects of involuntary eye movements on visual acuity. *Journal of the Optical Society of America*, *50*(8), 769-773.

Koenderink, J. J. (1972). Contrast enhancement and the negative afterimage. *Journal of the Optical Society of America*, *62*(5), 685-689.

Lemoine, L., Torre, K., & Delignieres, D. (2006). Testing for the presence of 1/f noise in continuation tapping. *Canadian Journal of Experimental Psychology, 60*, 247-

257.

Michaels, C. F., Arzamarski, R., Isenhower, R. W., & Jacobs, D. (2008). Direct learning in dynamic touch. *Journal of Experimental Psychology: Human Perception and Performance. 34*, 944-957.

Mill, J. S. (2002 [1848]). *A System of Logic: Ratiocinative and Inductive*. University Press of the Pacific.

Mitry, J. (1990) *The Aesthetics and Psychology of the Cinema.* Bloomington, IN: Indiana University Press.

Morales, C. J., & Kolaczyk, E. D. (2002). Wavelet-based multifractal analysis of human balance. *Annals of biomedical engineering*, *30*(4), 588-597.

Neri, P., Morrone, M.C., and Burr, D. (1998) Seeing biological motion. *Nature 395,* 894–896.

Nevado, A., Turiel, A., & Parga, N. (2000). Scene dependence of the non-gaussian scaling properties of natural images. *Network: Computation in Neural Systems*, *11*(2), 131-152.

Palatinus, Zs., Dixon, J. A., & Kelty-Stephen, D. G. (2012). Fractal fluctuations in quiet standing predict the use of mechanical information for haptic perception. ***Annals of Biomedical Engineering, Vol. 41(8),*** **1625–1634.**

Paulus, W. M., Straube, A. & Brandt, Th. (1984). Visual stabilization of posture: physiological stimulus characteristics and clinical aspects. *Brain*, *107*(4), 1143-1163

Pavlova, M., and Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics, 62,* 889–899.

Peng, C.-K., Havlin, S., Stanley, H. E., Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos, 5*, 82-87.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E, 49,* 1685-1689.

Ratliff, F., & Riggs, L. A. (1950). Involuntary motions of the eye during monocular fixation. *Journal of experimental psychology*, *40*(6), 687.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology.* New York: Oxford University Press.

Reid, T. (1785). *Essays on the Intellectual Powers of Man.*

Rhodes, T. & Turvey, M. T. T. (2007). Human memory retrieval as Lévy foraging. *PhysicaA,385*, 255-260.

Riggs, L. A., Ratliff, F., Cornsweet, J. C., & Cornsweet, T. N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America*, *43*(6), 495-500.

Rucci, M., & Desbordes, G. (2003). Contributions of fixational eye movements to the discrimination of briefly presented stimuli. *Journal of Vision*, *3*(11).

Runeson, S. (1988). The distorted room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception and Performance, 14* (2), 295-304.

Schertzer, D. & Lovejoy, S. (2004). Space-time Complexity and Multifractal Predictability, *Physica A*, ***338****,* 173-186.

Shlesinger, M. F., Zaslavsky, G. M. & Klafter, J.(1993). Strange kinetics. *Nature, 363,* 31-37

Shayeganfar, F., Jabbari-Farouji, S., Movahed, M. S., Jafari, G. R., & Tabar, M. R. R. (2009). Multifractal analysis of light scattering-intensity fluctuations. *Physical Review E, 80*(6), 061126.

Shelhamer, M., & Joiner, W. M. (2003). Saccades exhibit abrupt transition between reactive and predictive, predictive saccade sequences have long-term correlations. *Journal of Neurophysiology*, *90*(4), 2763-2769.

Shipley, T.F. (2003). The effect of object and event orientation on perception of biological motion. *Psychological Science 14,* 377–380.

Simion, F., Regolin, L. & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences, 105*(2), 809–813.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Steinman, R. M., Haddad, G. M., Skavenski, A. A., & Wyman, D. (1973). Miniature eye movement. *Science, 181*(4102), 810-819.

Stephen, D. G., Anastas, J. R., & Dixon, J. A. (2012). Scaling in executive control reflects multiplicative multifractal cascade dynamics. *Frontiers in Physiology-Fractal Physiology, 3,* 102.

Stephen, D. G., & Anastas, J. (2011). Fractal fluctuations in gaze speed visual search. *Attention, Perception, & Psychophysics, 73,* 666-677.

Stephen, D. G, & Dixon, J. A. (2009). The self-organization of insight: Entropy and power laws in problem solving. *Journal of Problem Solving, 2*, 72-101.

Stephen, D. G., & Hajnal, A. (2011). Transfer of calibration between hand and foot: Functional equivalence and fractal fluctuations. Attention, Perception, & Psychophysics, 73, 1302-1328.

Stephen, D. G., & Mirman, D. (2010). Interactions dominate the dynamics of visual cognition. *Cognition, 115,* 154-165.

Stephen, D. G., Boncoddo, R. A., Magnuson, J. S., & Dixon, J. A. (2009). The dynamics of insight: Mathematical discovery as a phase transition. *Memory & Cognition, 37*(8), 1132-1149.

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception 13,* 283–286.

Troje, N.F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision. 2*, 371– 387.

Tulunay-Keesey, U., & Jones, R. M. (1976). The effect of micromovements of the eye and exposure duration on contrast sensitivity. *Vision Research*, *16*(5), 481-488.

Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition, 9*, 237-304.

Vallortigara, G., Regolin, L., and Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biol. 3,* e208.

Valsiner, J. (2007). Gilbert Gottlieb’s Theory of Probalisitc Epigenesis: Probabilities and Realities in Development. *Developmental Psychobiology.* DOI 10.1002/dev.20276.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General, 132,* 331-350.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2005). Human cognition and 1/f scaling. *Journal of Experimental Psychology: General*, *134*, 117-123.

Webb, J. & Aggarwal, J. (1982). Structure from motion of rigid and jointed objects, *Artificial Intelligence 19*(1), 107–130

Withagen, R., & Michaels, C. F. (2005). Information for calibration and information for attunement in length perception by dynamic touch. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1379-1390.

Yarbus, A. L. (1967). *Eye movements and vision.* New York:Plenum Press.

Table 1.

Coefficients for model 1: Prediction of the event of recognition

	B	SE	*z*	*P*
(Intercept)	-3.7652	0.5974	-6.303	< .0001
epoch	0.2228	0.0277	8.035	< .0001
trial	0.1446	0.0334	4.328	< .0001
as.factor(camera)	-0.1058	0.1704	-0.621	0.5348
as.factor(PLD)2	1.7687	0.4104	4.309	< .0001
as.factor(PLD)3	1.2561	0.3910	3.213	0.0013
as.factor(PLD)4	1.6343	0.4045	4.041	< .0001
as.factor(PLD)5	0.1666	0.3865	0.431	0.6664
as.factor(PLD)6	-0.3479	0.3881	-0.897	0.3699
as.factor(PLD)7	1.1436	0.3857	2.965	0.0030
as.factor(PLD)8	1.9829	0.4059	4.886	< .0001
as.factor(PLD)9	1.0659	0.3869	2.755	0.0059
as.factor(PLD)10	1.5244	0.3999	3.812	< .0001
MF spectrum width	-0.7951	0.1457	-5.458	< .0001

Table 2.

Coefficients for model 2: Prediction of the event of recognition

	B	SE	*z*	*P*
(Intercept)	-4.8730	0.8616	-5.655	< .0001
epoch	0.1973	0.0352	5.612	< .0001
as.factor(trial)2	0.4795	0.8019	0.598	0.54983
as.factor(trial)3	-0.6710	0.9230	-0.727	0.46723
as.factor(trial)4	-1.1400	0.9679	-1.177	0.23908
as.factor(trial)5	-0.1260	0.9483	-0.133	0.89434
as.factor(trial)6	1.0020	0.8343	1.201	0.22984
as.factor(trial)7	0.0257	0.8577	0.03	0.97614
as.factor(trial)8	0.3692	0.8625	0.428	0.6686
as.factor(trial)9	-0.4621	0.9047	-0.511	0.60947
as.factor(trial)10	0.7955	0.7693	1.034	0.30107
as.factor(PLD)2	0.2975	0.6276	0.474	0.63547
as.factor(PLD)3	-0.3532	0.7502	-0.471	0.63773
as.factor(PLD)4	1.8050	0.6499	2.777	0.00548
as.factor(PLD)5	-1.0570	0.7340	-1.44	0.14995
as.factor(PLD)6	0.0023	0.6570	0.003	0.99722
as.factor(PLD)7	0.0431	0.7121	0.061	0.95173
as.factor(PLD)8	1.2910	0.1900	0.001	0.99946
as.factor(PLDtype)2	1.0860	0.4729	2.296	0.02167
as.factor(PLDtype)3	0.8740	0.4446	1.966	0.04934
as.factor(scrambled)1	-19.2400	1.0660	-0.018	0.9856
MF spectrum width	-0.7461	0.1639	-4.551	< .0001

Table 3.

Coefficients for model 3: Prediction of the event of recognition

	B	SE	*z*	*P*
(Intercept)	-5.1596	0.4869	-10.597	< .0001
epoch	0.1355	0.0222	6.106	< .0001
trial	-0.0789	0.0503	-1.569	0.1165
as.factor(PLD)2	0.2936	0.2627	1.118	0.2637
as.factor(PLD)3	-1.3551	0.3463	-3.913	< .0001
as.factor(PLD)4	-0.7244	0.3024	-2.396	0.0166
as.factor(PLD)5	-1.3232	0.3482	-3.8	< .0001
as.factor(PLDtype)2	1.0274	0.3029	3.392	< .0001
as.factor(PLDtype)3	1.6559	0.2936	5.64	< .0001
as.factor(exposure)2	1.3030	0.3748	3.477	< .0001
as.factor(exposure)3	1.8607	0.5870	3.17	0.0015
MF spectrum width	-0.2825	0.0878	-3.217	0.0013

Table 4.

Coefficients for model 4: Prediction of the event of recognition

	B	SE	*z*	*P*
(Intercept)	-6.4743	0.7758	-8.3460	< .0001
epoch	0.1721	0.0305	5.6480	< .0001
trial	0.2272	0.0642	3.5380	< .001
as.factor(PLD)2	0.4113	0.5981	0.6880	0.4916
as.factor(PLD)3	-1.5987	0.8230	-1.9430	0.0521
as.factor(PLD)4	3.1636	0.4895	6.4620	< .0001
as.factor(PLD)5	2.7464	0.4389	6.2570	< .0001
as.factor(PLD)6	3.3400	0.4435	7.5310	< .0001
as.factor(PLD)7	-1.0778	0.4707	-2.2900	0.0220
as.factor(PLD)8	-1.4356	0.5634	-2.5480	0.0108
as.factor(PLD)9	-1.3944	0.4879	-2.8580	0.0043
as.factor(PLDtype)2	0.5604	0.2311	2.4250	0.0153
as.factor(congruency)1	0.8102	0.2407	3.3660	< .001
MF spectrum width	-0.4767	0.1036	-4.6020	< .0001

Table 5.

Coefficients for model 5: Prediction of the event of recognition

	B	SE	*z*	*P*
(Intercept)	-4.2667	0.5530	-7.7150	< .0001
epoch	0.1605	0.0275	5.8310	< .0001
trial	0.0542	0.0413	1.3130	0.1893
as.factor(PLD)2	0.6409	0.4499	1.4240	0.1543
as.factor(PLD)3	-16.3086	582.4260	-0.0280	0.9777
as.factor(PLD)4	1.5022	0.4439	3.3840	< .001
as.factor(PLD)5	3.0527	0.4820	6.3340	< .0001
as.factor(PLD)6	2.9747	0.4771	6.2360	< .0001
as.factor(PLD)7	-1.8416	0.7864	-2.3420	0.0192
as.factor(PLD)8	-2.5601	1.0578	-2.4200	0.0155
as.factor(PLD)9	-0.5428	0.5351	-1.0140	0.3104
as.factor(PLD)10	1.4725	0.4313	3.4140	< .001
as.factor(cam)1	NA	NA	NA	NA
MF spectrum width	-0.5226	0.1154	-4.5270	< .0001

Table 6.

PLD animations in Experiment 1 and 2a

code	name	action
a	avoidance	actor swatting at feet
b	ballet	ballerina twirl and leap
c	catch ball	catching and throwing a ball
d	shoveling	digging a hole in the ground
f	fishing	casting and reeling fishing line
g	getting up	standing up from the ground
k	kick	roundhouse kick
l	limping	walking with a limp
r	imitate ape	actor imitates the arm movements of an ape
s	sweeping	actor sweeps the floor with a broom

Figure captions

Figure 1. Gottlieb's model of Probabilistic Epigenesis. Adopted from Gottlieb (2007).

Figure 2. Expansion of Gottlieb's model taylored to visual perception. The original levels are replaced by levels of the visual perceptual system and the visual environment.

Figure 3. Experiment 1, Interaction plot of mean response times by PLD animations and camera conditions.

Figure 4. Experiment 2a, Interaction plot of mean response times by PLD animations and camera conditions.

Figure 5. Experiment 2a, Survival functions for still and moving conditions. Proportion of participants still before recognition (y axis) is plotted against time, represented by epochs (x axis). The rate of dropping out from the risk set does not differ significantly in the two conditions.

Figure 6. Experiment 2a, Hazard functions (odds) for still and moving conditions.

Figure 7. Experiment 2a, Surface plot of survival functions and MF width quartiles. The y axis represents the proportion of participants at risk as a function of epochs (x axis) and MF width quartiles (z axis). Survival functions for the four width quartiles are represented by the lines (wider is further away from the origo), the effect size is indexed by the size of the segments connecting the lines. Color represents the proportion of participants at risk (before recognition).

Figure 8. Experiment 2b, Plot of proportion of correct responses by camera conditions.

Figure 9a. Experiment 2b, Survival functions for the three camera conditions.

Figure 9b. Experiment 2b, Survival functions for the normal and scrambled conditions.

Figure 10. Experiment 2b, Surface plot of survival functions and MF width quartiles.

Figure 11. Experiment 3a, Plot of proportion of correct responses by camera conditions.

Figure 12. Experiment 3a, Plot of proportion of correct responses by exposure.

Figure 13a. Experiment 3a, Survival functions for the three camera conditions.

Figure 13b. Experiment 3a, Survival functions for the three exposure conditions.

Figure 14. Experiment 3a, Surface plot of survival functions and MF width quartiles.

Figure 15a. Experiment 3b, Plot of proportion of correct responses by camera conditions.

Figure 15b. Experiment 3b, Plot of proportion of correct responses by congruency.

Figure 16a. Experiment 3b, Survival functions for the three camera conditions.

Figure 16b. Experiment 3b, Survival functions for the congruency.

Figure 17. Experiment 3b, Surface plot of survival functions and MF width quartiles.

Figure 18a. Experiment 3c, Plot of percentage of correct responses by camera conditions.

Figure 18b. Experiment 3c, Plot of percentage of correct responses by switch in context.

Figure 19a. Experiment 3c, Survival functions for the two camera conditions.

Figure 19b. Experiment 3c, Survival functions for the switch in context.

Figure 20. Experiment 3c, Surface plot of survival functions and MF width quartiles.

Figure 1.

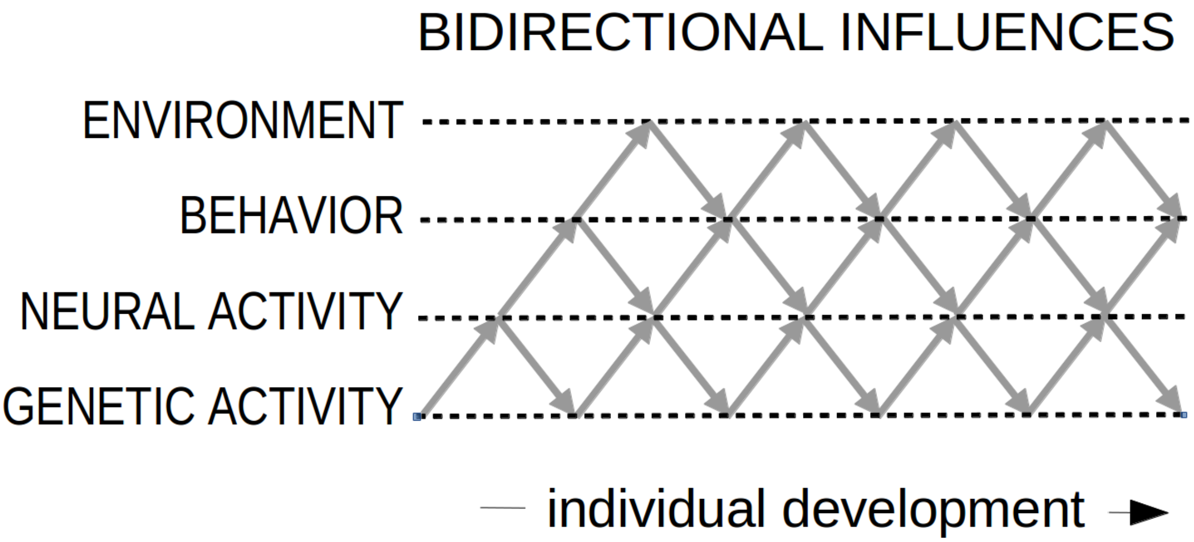

BIDIRECTIONAL INFLUENCES
ENVIRONMENT
BEHAVIOR
NEURAL ACTIVITY
GENETIC ACTIVITY
individual development

Figure 2.

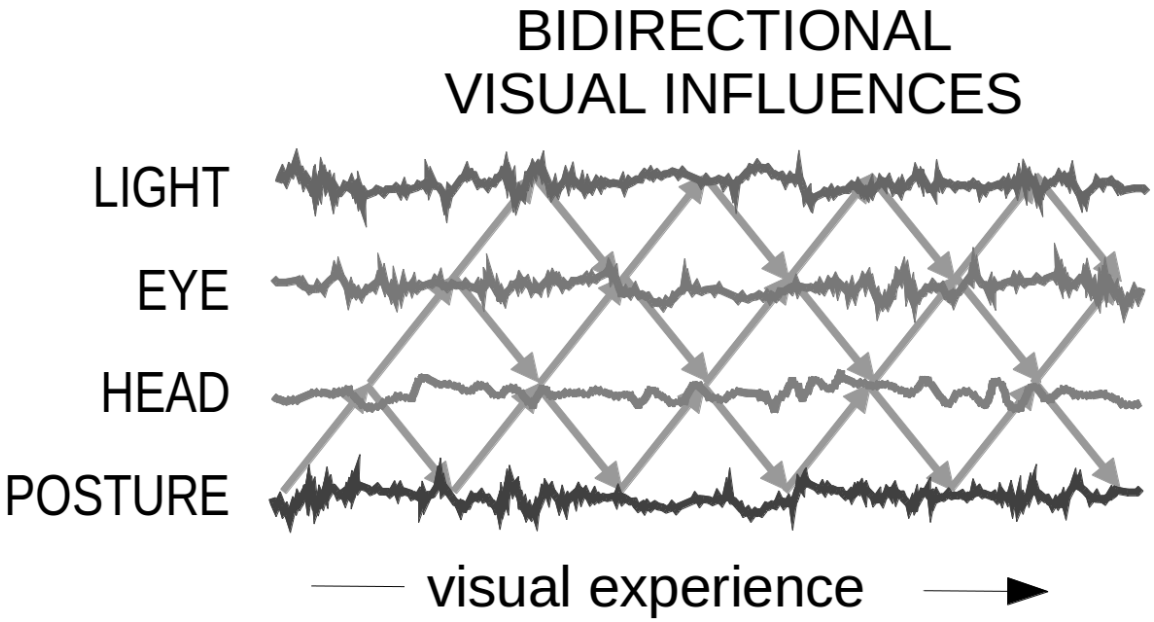

BIDIRECTIONAL
VISUAL INFLUENCES
LIGHT
EYE
HEAD
POSTURE
visual experience

Figure 3.

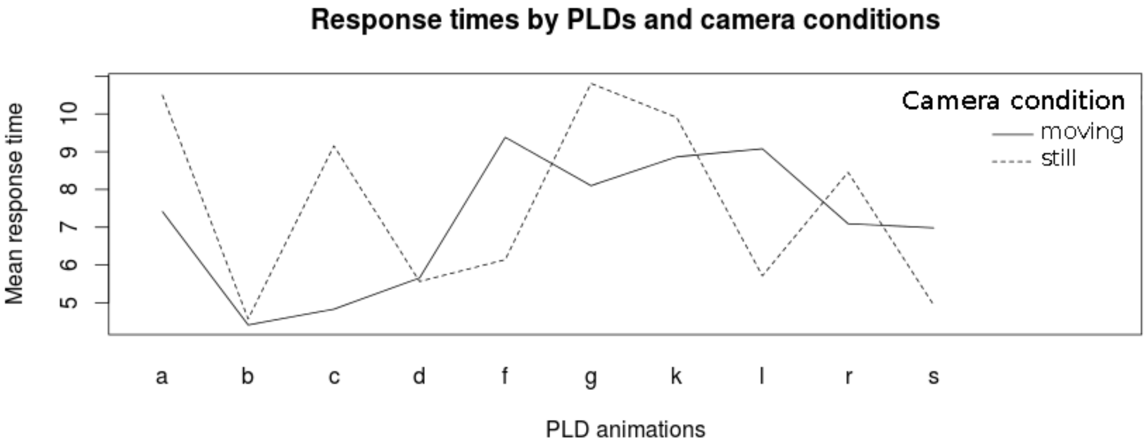

Response times by PLDs and camera conditions
Mean response time
5
6
7
8
9
10
a
b
c
d
f
g
k
l
r
s
PLD animations
Camera condition
moving
still

Figure 4.

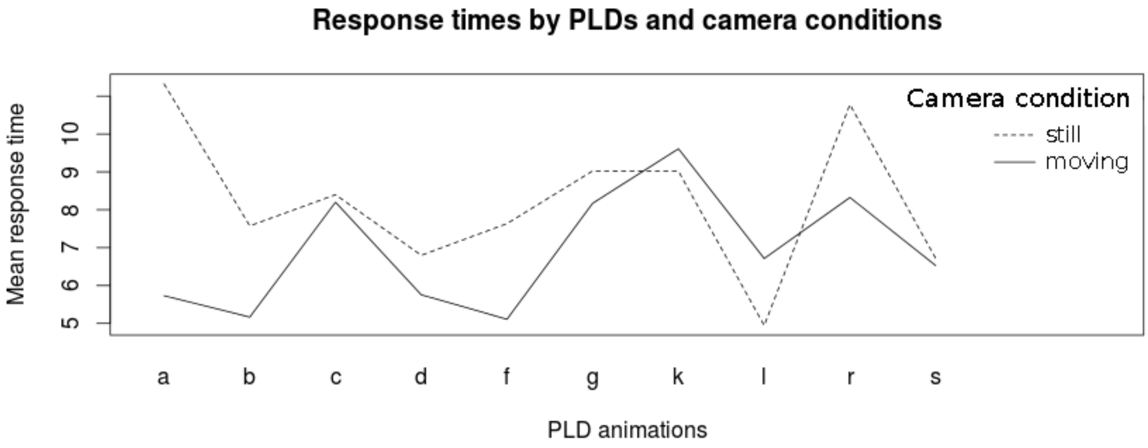
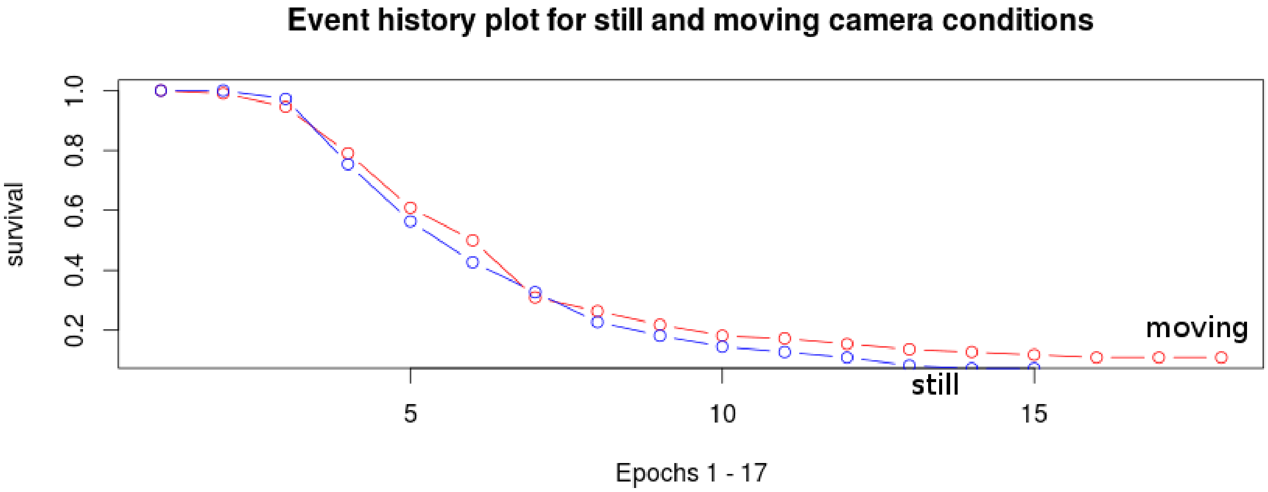
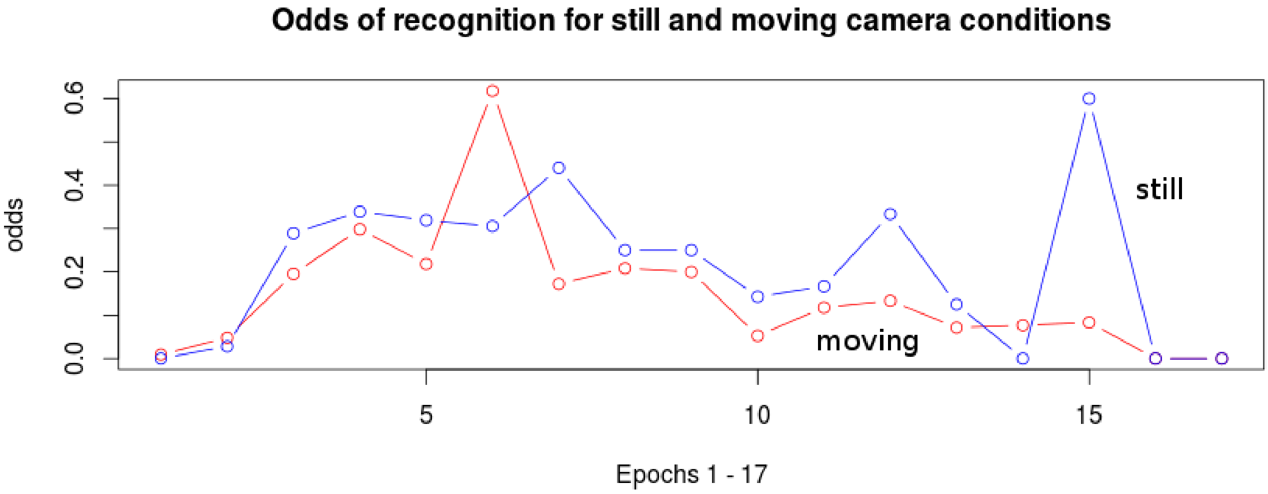

Response times by PLDs and camera conditions
Mean response time
5
6
7
8
9
10
Camera condition
still
moving
a
b
c
d
f
g
k
l
r
s
PLD animations

Figure 5



Event history plot for still and moving camera conditions
survival
0.2
0.4
0.6
0.8
1.0
5
10
15
Epochs 1 - 17
moving
still

Figure 6.



Odds of recognition for still and moving camera conditions
odds
0.0
0.2
0.4
0.6
5
10
15
Epochs 1 - 17
still
moving

Figure 7

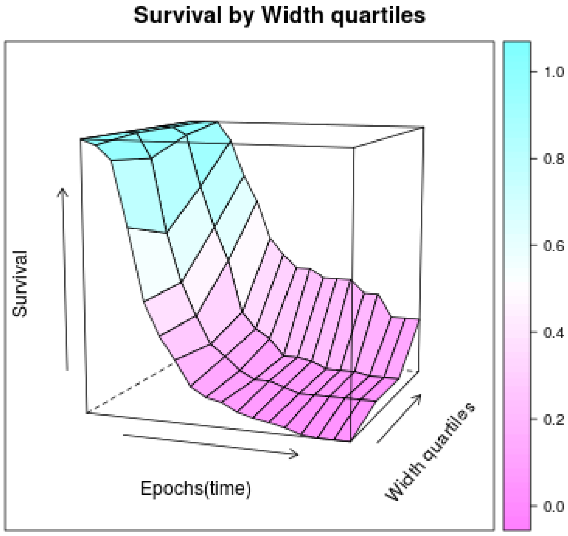

Survival by Width quartiles
Survival
Epochs(time)
Width quartiles
1.0
0.8
0.6
0.4
0.2
0.0

Figure 8.

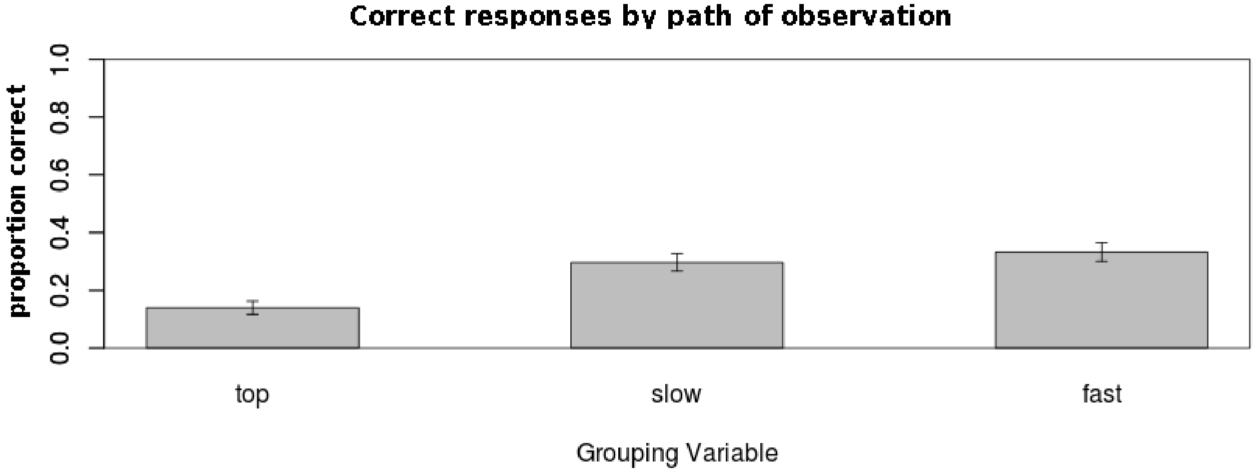
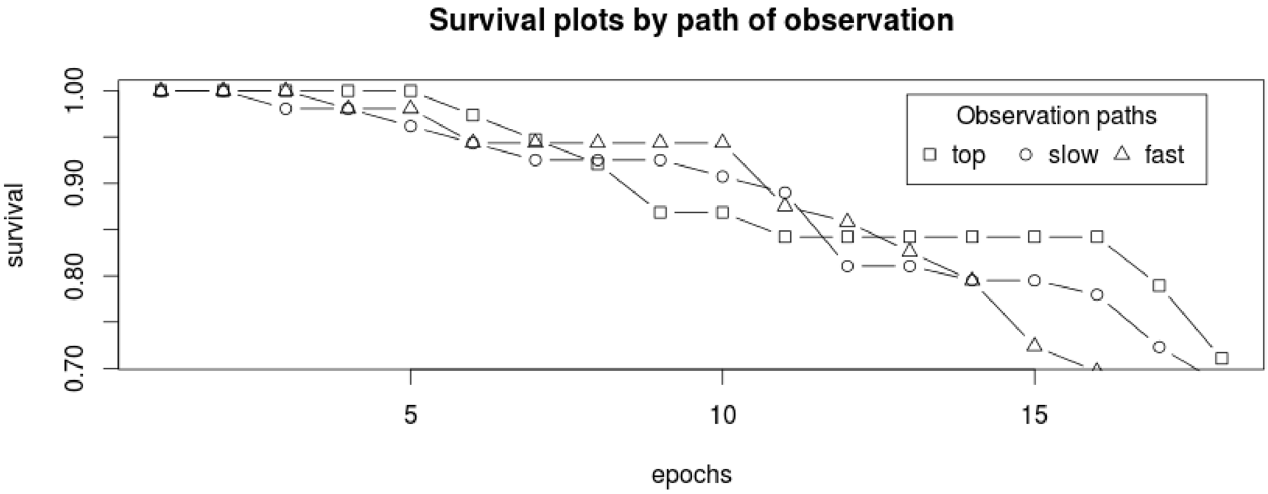

Correct responses by path of observation
proportion correct
0.0
0.2
0.4
0.6
0.8
1.0
top
slow
fast
Grouping Variable

Figure 9

a


Survival plots by path of observation
survival
1.00
0.90
0.80
0.70
epochs
5
10
15
Observation paths
□ top
○ slow
△ fast

b

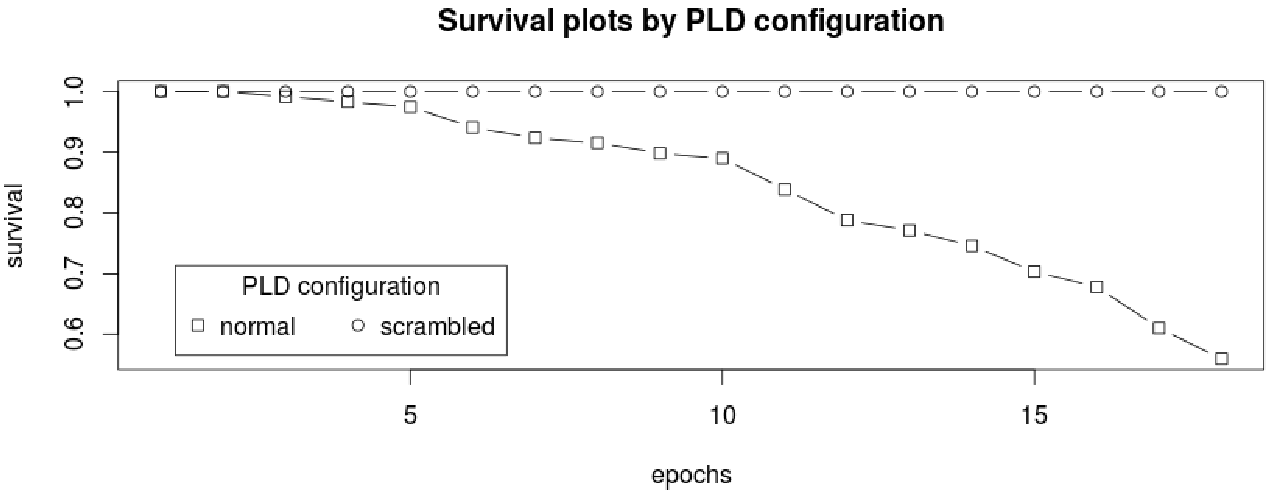
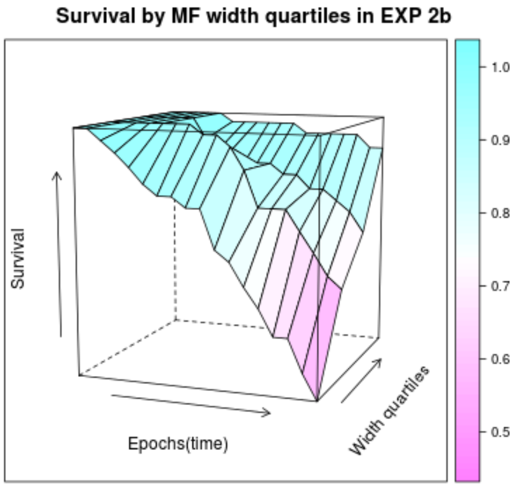
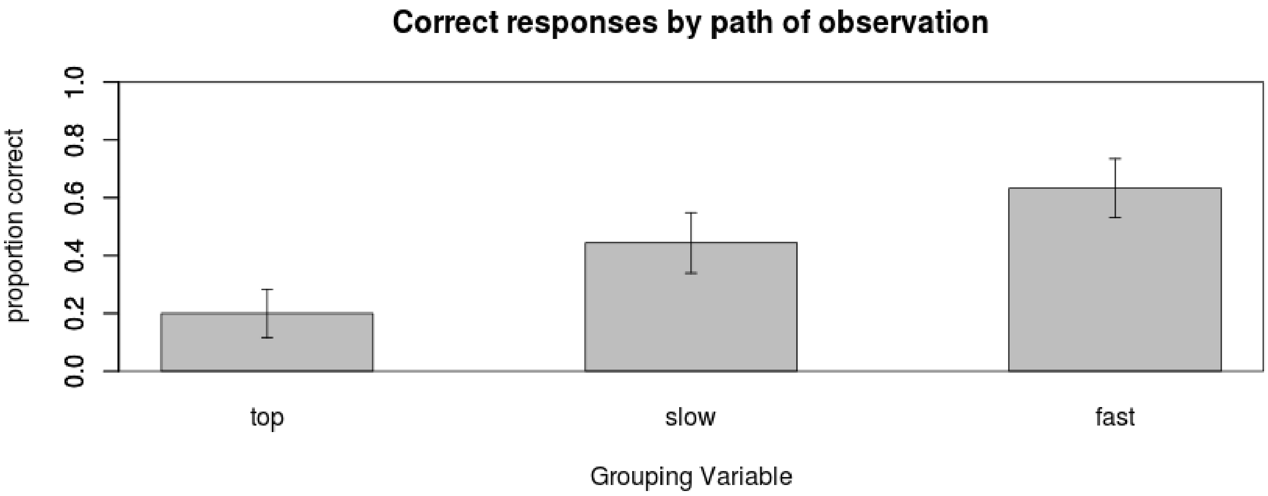

Survival plots by PLD configuration
survival
1.0
0.9
0.8
0.7
0.6
epochs
5
10
15
PLD configuration
□ normal
○ scrambled

Figure 10.



Survival by MF width quartiles in EXP 2b
Survival
Epochs(time)
Width quartiles
1.0
0.9
0.8
0.7
0.6
0.5

Figure 11.



Correct responses by path of observation
proportion correct
0.0
0.2
0.4
0.6
0.8
1.0
top
slow
fast
Grouping Variable

Figure 12.

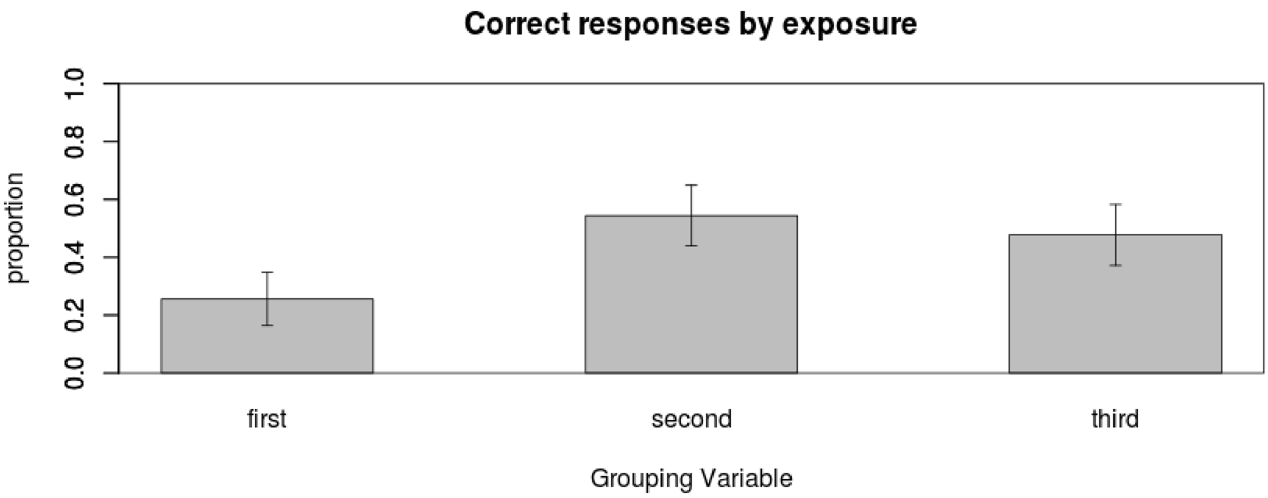
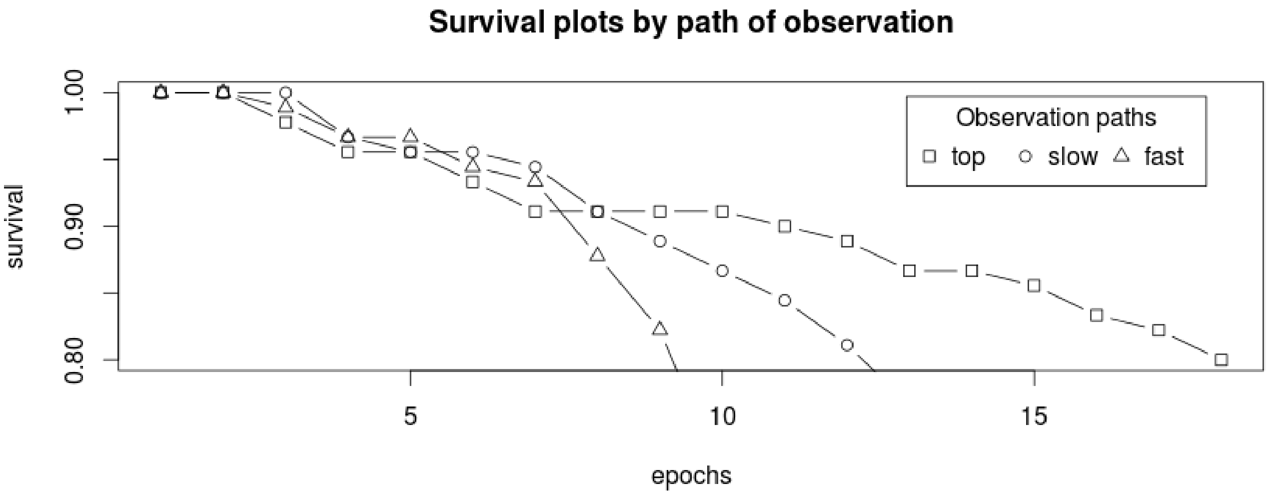

Correct responses by exposure
proportion
0.0
0.2
0.4
0.6
0.8
1.0
first
second
third
Grouping Variable

Figure 13.

a


Survival plots by path of observation
survival
1.00
0.90
0.80
epochs
5
10
15
Observation paths
□ top
○ slow
△ fast

b

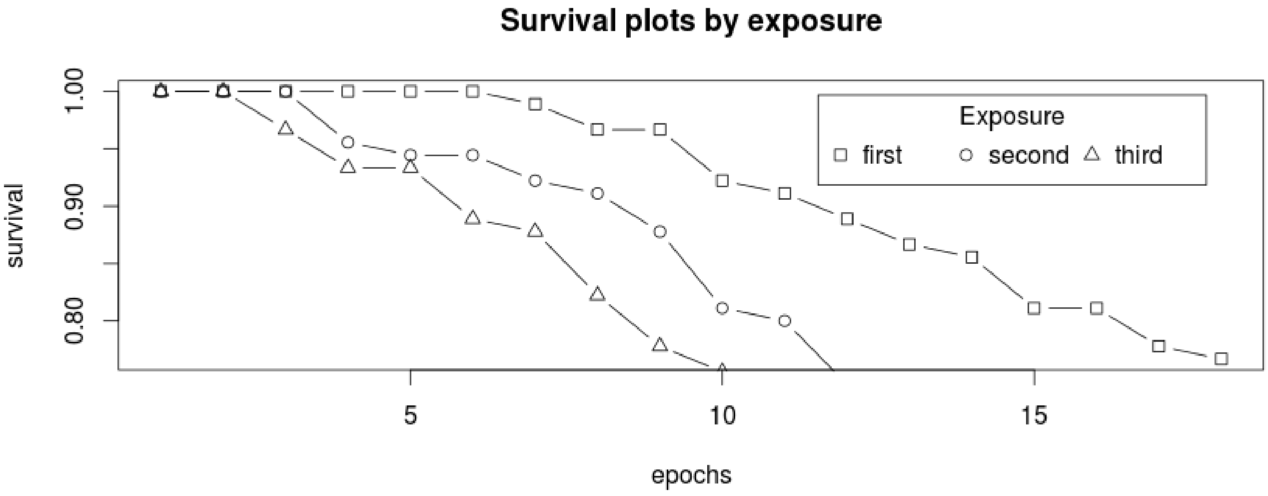

Survival plots by exposure
survival
1.00
0.90
0.80
epochs
5
10
15
Exposure
□ first
○ second
△ third

Figure 14.

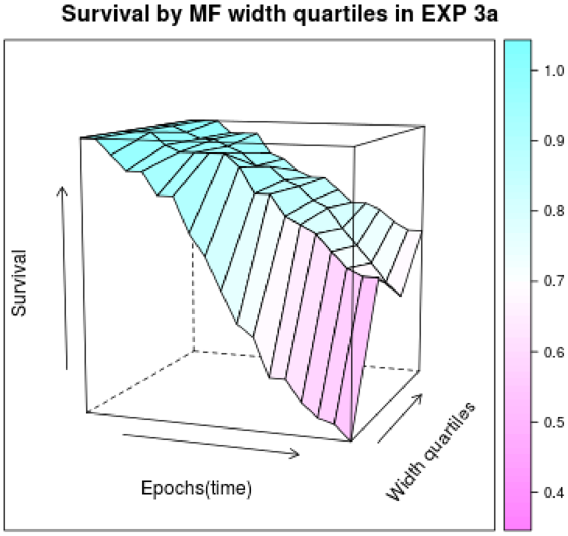

Survival by MF width quartiles in EXP 3a
Survival
Epochs(time)
Width quartiles
1.0
0.9
0.8
0.7
0.6
0.5
0.4

Figure 15.

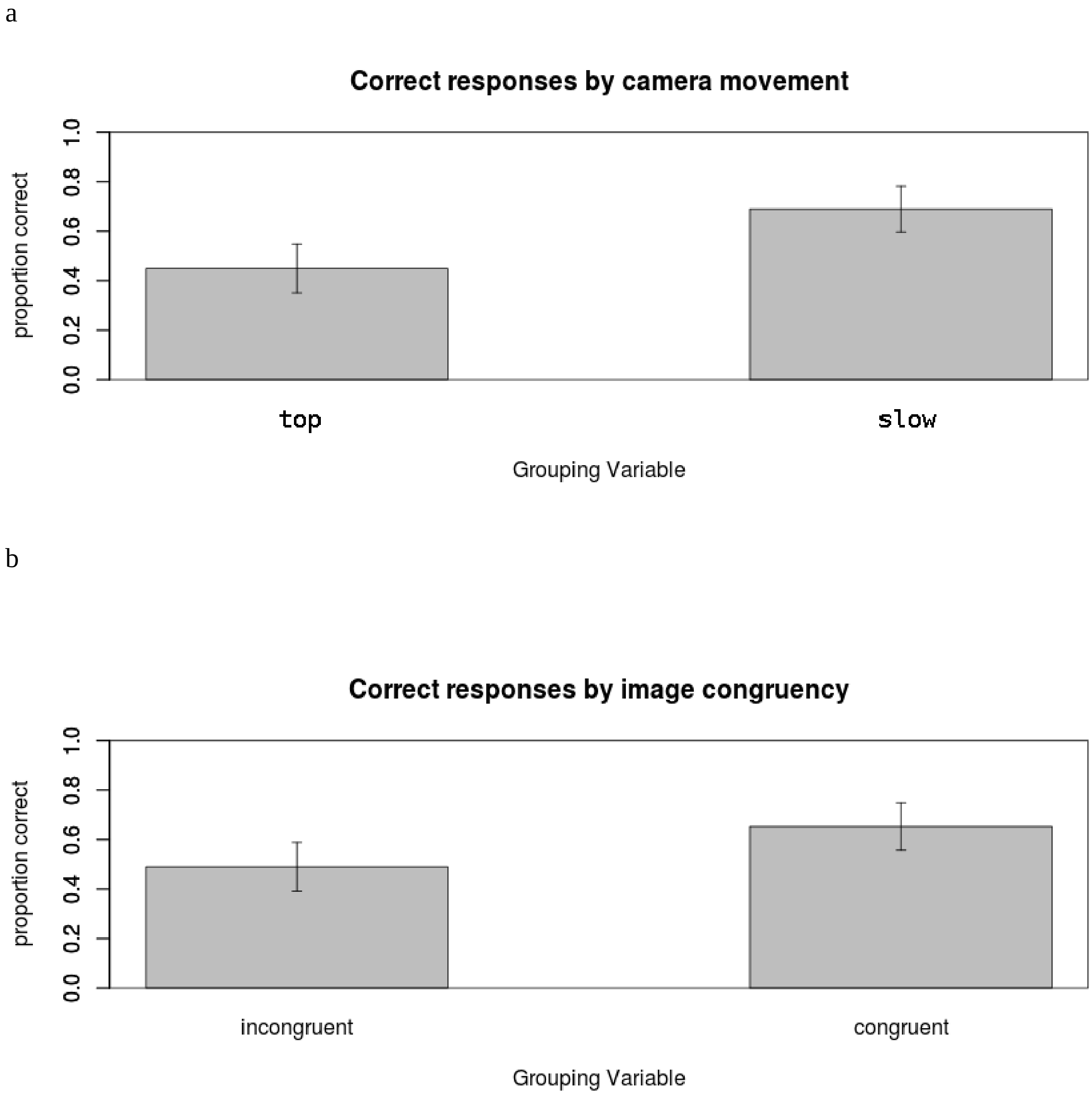

a
Correct responses by camera movement
proportion correct
0.0
0.2
0.4
0.6
0.8
1.0
top
slow
Grouping Variable
b
Correct responses by image congruency
proportion correct
0.0
0.2
0.4
0.6
0.8
1.0
incongruent
congruent
Grouping Variable

Figure 16.

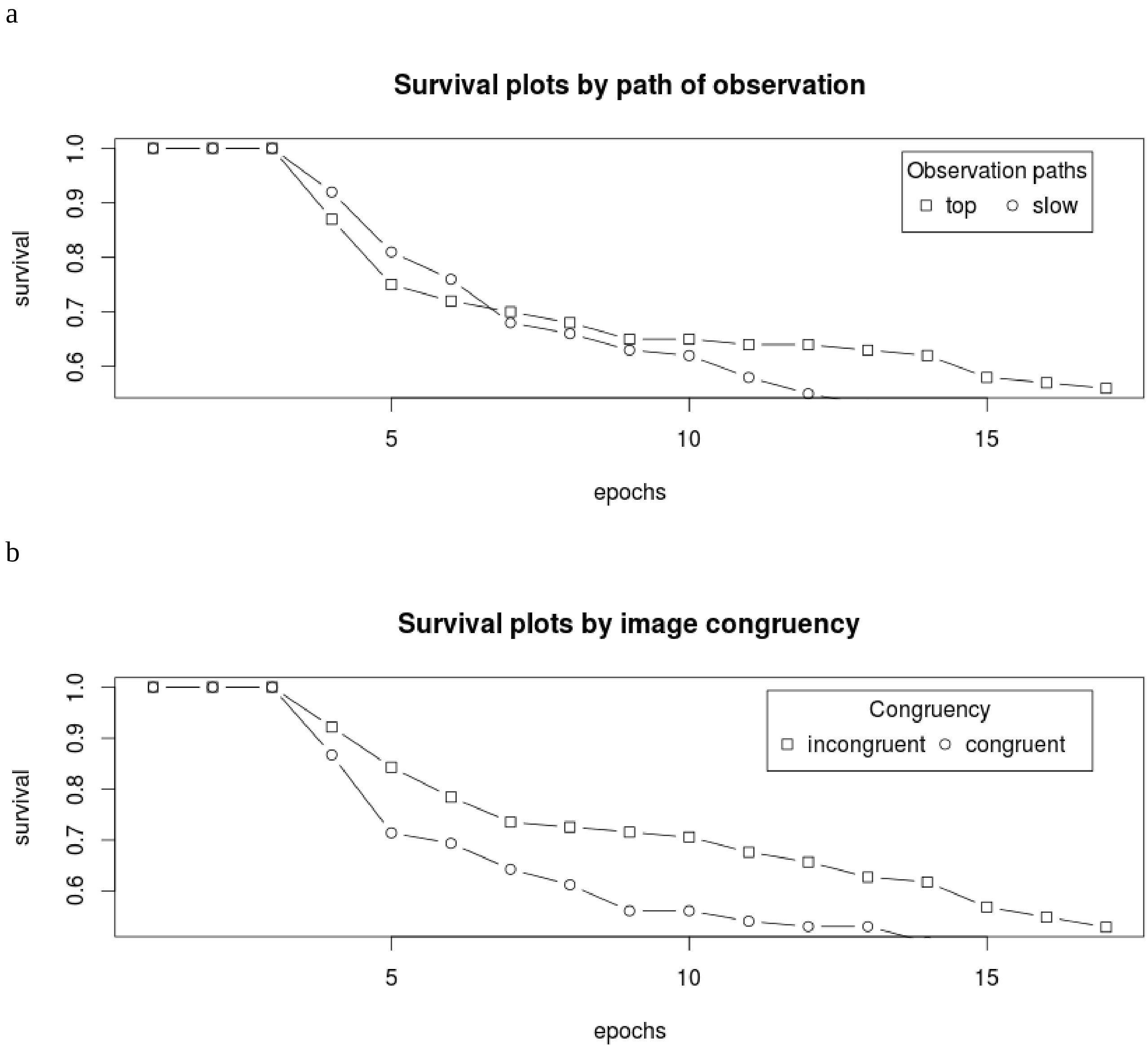

a
Survival plots by path of observation
Observation paths
top
slow
survival
0.6
0.7
0.8
0.9
1.0
5
10
15
epochs
b
Survival plots by image congruency
Congruency
incongruent
congruent
survival
0.6
0.7
0.8
0.9
1.0
5
10
15
epochs

Figure 17.

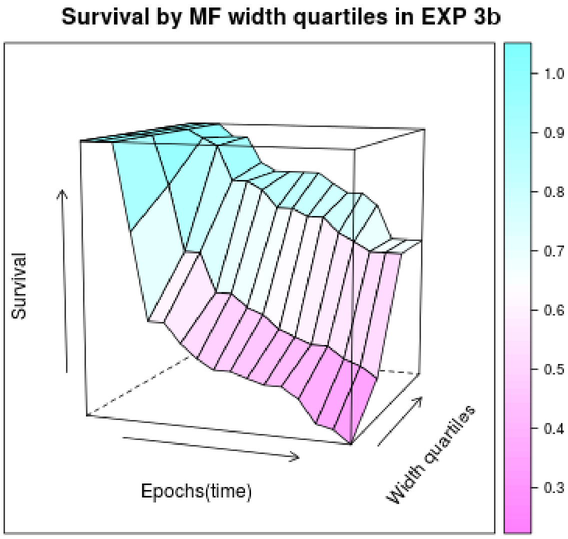

Survival by MF width quartiles in EXP 3b
Survival
Epochs(time)
Width quartiles
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3

Figure 18.

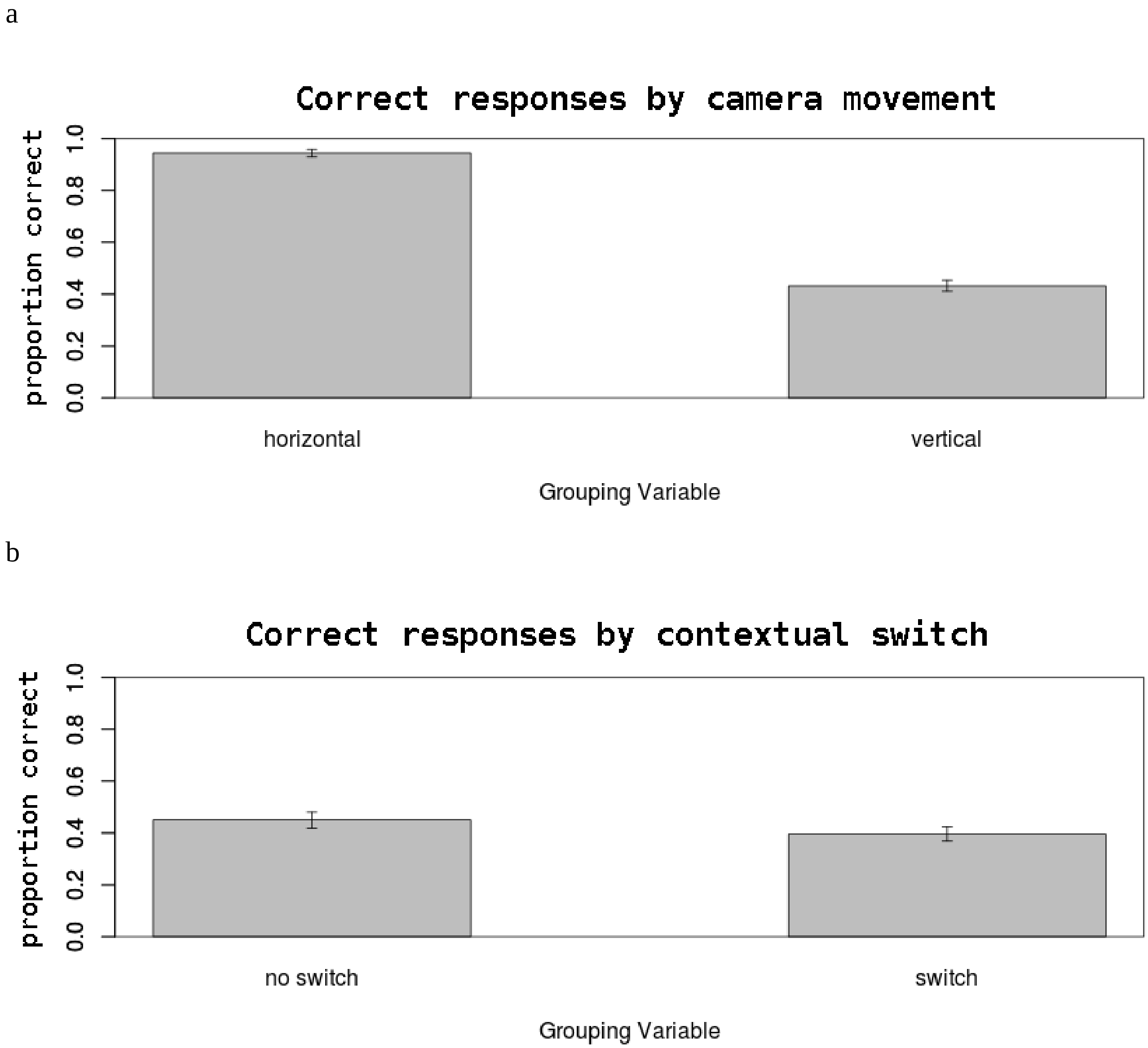

a
Correct responses by camera movement
proportion correct
0.0
0.2
0.4
0.6
0.8
1.0
horizontal
vertical
Grouping Variable
b
Correct responses by contextual switch
proportion correct
0.0
0.2
0.4
0.6
0.8
1.0
no switch
switch
Grouping Variable

Figure 19.

a

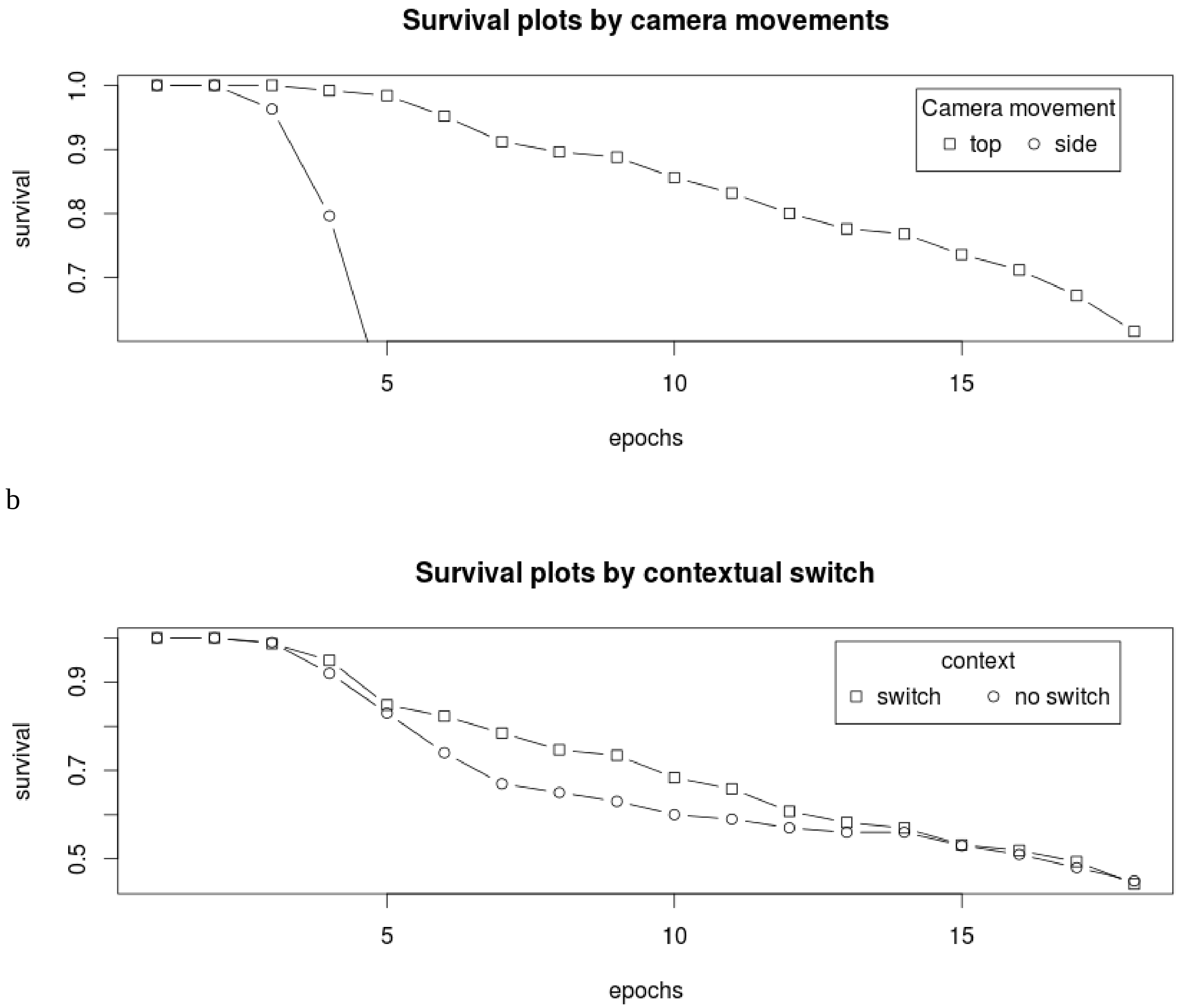
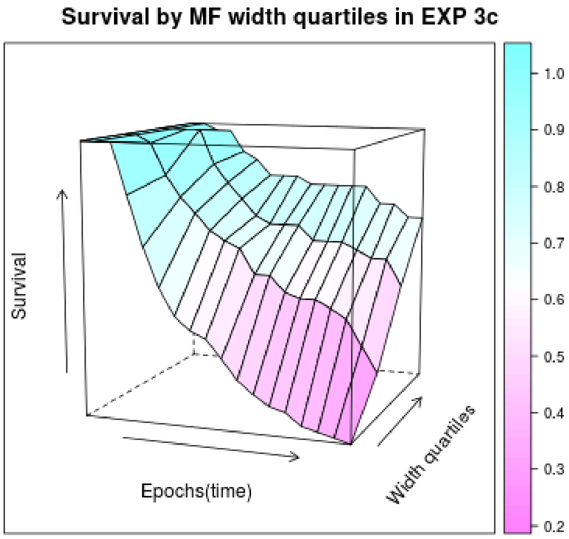

Survival plots by camera movements
survival
1.0
0.9
0.8
0.7
Camera movement
□ top
○ side
5
10
15
epochs
b
Survival plots by contextual switch
survival
0.9
0.7
0.5
context
□ switch
○ no switch
5
10
15
epochs

Figure 20.


Survival by MF width quartiles in EXP 3c
Survival
Epochs(time)
Width quartiles
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2