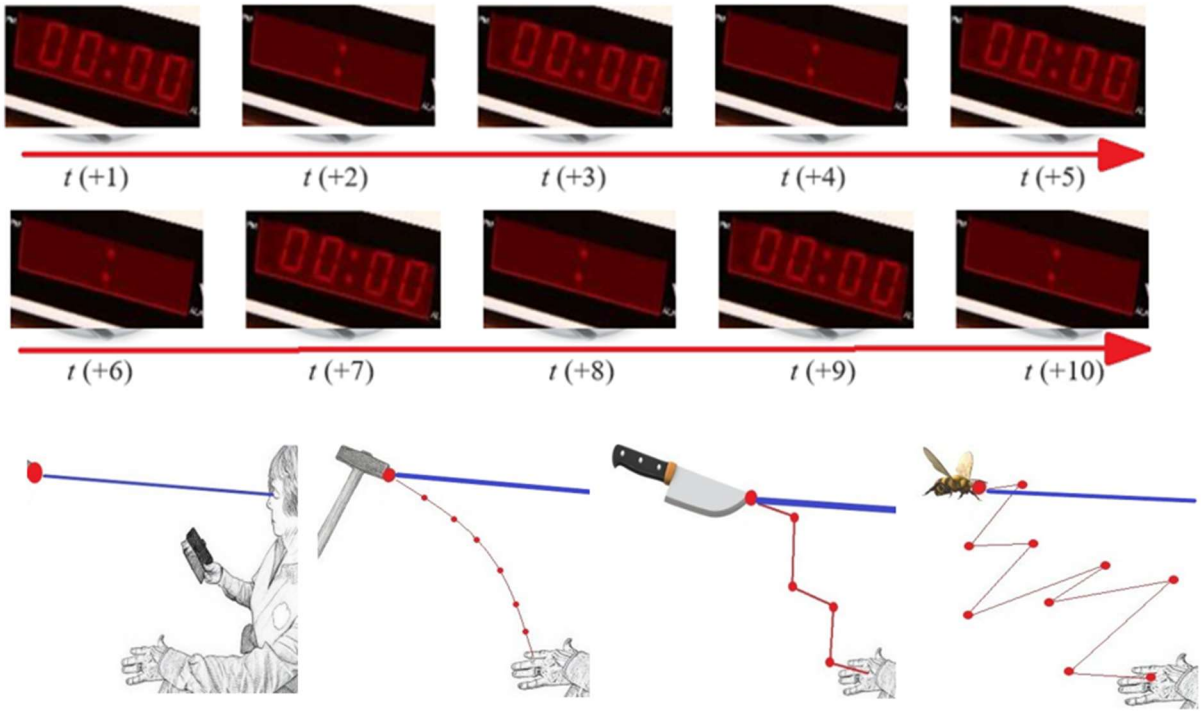
The Origin of Consciousness

Attentional Functions as Continuous Action-Oriented Comparison Processes



Caught In A Line
The explanatory model of all motoric movement actions

N.J. Mol
December 2025 ©

Contact: kwillinq@gmail.com
https://www.researchgate.net/profile/Nj-Mol/research
https://www.explanatorymodel.nl/

Abstract

Current vision science still lacks a coherent mechanistic framework, largely because perception and consciousness are often treated as representational or intermittently activated processes rather than as continuously operating control functions. This article presents an ecologically grounded architecture in which auditory and visual perception originate from a single primordial survival requirement: the continuous detection and evaluation of environmental change in order to regulate action under dynamic and potentially conflicting conditions.

Perception is therefore not capacity-limited or episodic, but an always-on comparison process that evaluates differences across adjacent temporal frames. Consciousness is identified as the direct experiential expression of this evaluative comparator activity, emerging precisely when perceptual information becomes relevant for ongoing or imminent action. In this view, consciousness is not an added cognitive layer, but an inherent property of a biological control system optimized for action selection.

Within this architecture, organisms respond to approaching objects through one auditory and four visual phases of increasing specificity. These phases progressively constrain possible action trajectories by refining estimates of movement, intersection likelihood, action trajectory shape, and *tau* (time-to-contact). Although described sequentially, all phases remain continuously active as parallel comparator streams. This multilayered structure reframes consciousness as the phenomenal manifestation of a continuously operating ecological control system dedicated to detecting, evaluating, and regulating movement in relation to the organism's own action possibilities.

Keywords: *Visual attention; auditory deviation detection; motion comparison mechanism; adjacent temporal frames; ecological perception; action trajectory shape; tau (time-to-contact); autonomous perceptual phases; deviation detection; focal vision; peripheral vision; continuous baseline comparator; predictive updating; egocentric movement preparation.*

Introduction

Despite decades of empirical research, vision science and consciousness research still lack a unified mechanistic account of how perceptual processes support adaptive action. This persistent gap can be traced to a series of categorical errors: perception is commonly conceptualized as a representational input system, attention as a selectively activated resource, and consciousness as an optional or emergent add-on rather than as a core biological function. As a result, existing frameworks struggle to

explain how organisms reliably select and regulate actions under continuously changing and potentially conflicting environmental conditions.

This article approaches consciousness from a different starting point. Rather than asking how perceptual representations give rise to subjective experience, it begins with the evolutionary problem of action regulation. For any organism, survival depends on the ability to detect environmental change early enough to evaluate its relevance for ongoing or future action. This requirement precedes object recognition, categorization, or introspection and applies across species and levels of complexity.

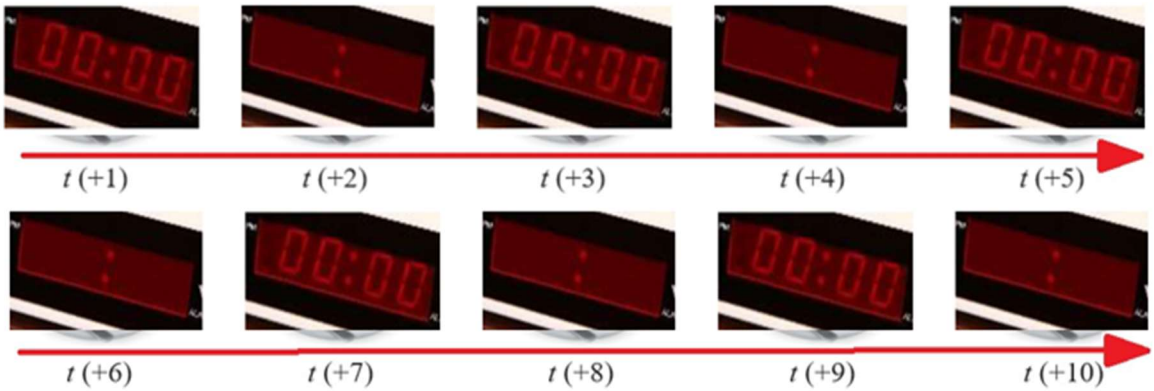
The explanatory model of motoric movement action provides a mechanistic solution to this problem. It shows that both auditory and visual perception arise from a single, continuously operating comparison mechanism that relates successive temporal frames in order to detect deviation, project action trajectories, and constrain possible responses. Conscious experience emerges precisely at this evaluative interface, where perceptual change becomes action-relevant. Consciousness is thus not explained as a by-product of complex representation, but as a functional property of an ecological control system optimized for regulating behavior under dynamic conditions.

Perception Organs Are Mainly Comparison Organs

The entire current discourse on conscious visual (and also auditory) perception overlooks a fundamental issue. Long before organisms developed any interest in what exactly was moving, they were solely interested in if something visually or auditorily moved. In order to establish contact (for feeding or mating) or to avoid contact (being eaten) of that environmental movement with their own egocentric movement which completely corresponds with the fight-or-flight narrative. For all organisms, the perception of pure motion in the environment—relative to the self (as described by Gibson)—holds the highest ecological and evolutionary value. Initially, we do not need to know what is moving, only that something is moving (potentially toward us).

This principle becomes evident in the example of a digital clock after a power outage. When the zeros on the display begin blinking, the perception of the first set of zeros have absolutely no relationship with subsequent sets of zeros. They only gain meaning through comparison.

This provides definitive scientific evidence that our perceptual organs/systems are, above all, comparison systems. Such comparison systems operate continuously; they do not require activation and therefore impose no reaction-time limitations on their earliest perceptual stages. Even more crucial is the implication derived from Gibson's notion of direct perception: motion is perceived consciously the very moment our eyes open.



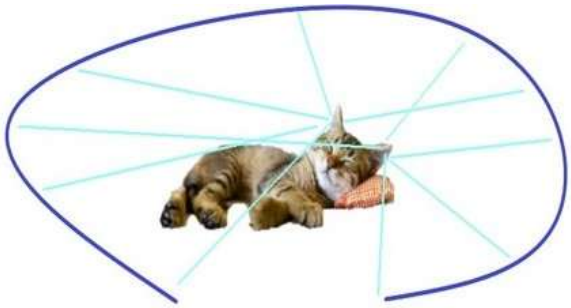
This requires a major paradigm shift in the science of perception: 1. Visual perception involves multiple autonomous phenomena, 2. All perceptual organs and systems are, in their essence, comparison systems—continuously relating separate, successive temporal images (as in a flipbook). In other words, all perceptual systems are fundamentally designed to, first and foremost, consciously detect movement, and 3. Zero-motion, as with stationary objects (e.g., the blinking zeros of a digital clock), is perceived just as actively as actual motion.

Crucially, the perception of the blinking zeros remains a perception of their motion, regardless of any later recognition that the shapes represent numerical digits. Object recognition forms a fully autonomous layer that is added on top of—rather than replacing—the ongoing motion-based comparison. Thus, the perceptual system maintains both processes in parallel: the continuous detection of change and the later, independent identification of what the changing pattern happens to be.

The Complete Ecological Account of Attention

The complete ecological explanation of all attentional foci encompasses one auditory and four visual phases of increasing specificity. Audition provides an initial 360° deviation detector. Visual phase 1 registers the first divergent point in the vista; phase 2 determines whether it reflects genuine motion and whether its latent continuation may intersect with the organism. Phase 3 identifies the object and uses its action trajectory shape and *tau*-value to project its continuation. Phase 4 refines this projection for time-critical motor preparation.

I. Auditory Phase



1

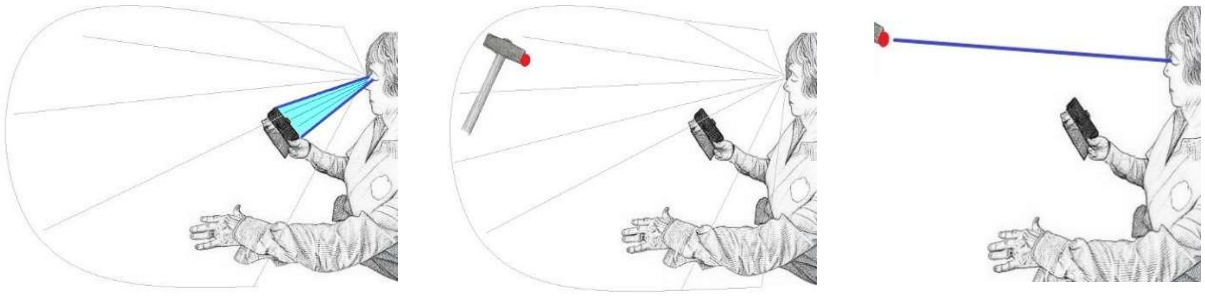
2

The primary level of motion detection encompasses the auditory attentional processes. They cover a 360° auditory field, which marks the lateral positioning of the ears—situated at the sides of the head, in relation to the eyes, close to the brain, and always active (even during sleep). During sleep (fig. 1) they remain functioning, albeit in a more basal form, which is far more parsimonious than keeping both eyes and ears fully engaged at all times. Note the by far strongest level of reduction (fig. 2): from the auditory detection phase of the ears (blue) to the peripheral visual attention phase of the eyes (red).

II. Visual Attention Phase 1a

Goal: Reduce peripheral visual attention from the entire vista to one point c.q. an extremely restricted region[1].

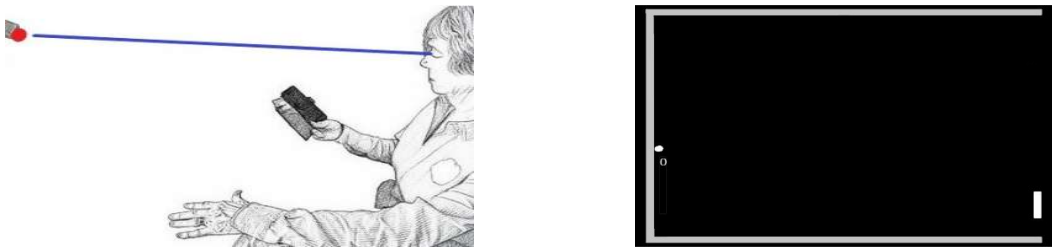
[1] The ecological primary objective within an optimization process.



1 2 3

Illustrations: *Within the Rubber Hand Illusion*[2] *a hammer suddenly appears within the visual field attacking the (rubber) hand.*

While direct gaze is directed at a book or a phone (fig.1), peripheral vision remains continuously attentive. This peripheral visual attention does not have to be activated; it is continuously present whenever the eyes are open and forms the baseline condition that enables phase 1a to occur. Within this remaining vista, the first level of visual motion detection concerns only whether *any* significant deviation occurs. At this stage, the system registers merely a detectable difference across adjacent temporal frames within the peripherally covered field. It is therefore *not* yet concerned with whether the stimulus is a hammer, nor whether it is moving. Figure 2 is therefore not an accurate representation of this first level of visual motion detection; Figure 3 illustrates it much better by depicting the hammerhead as an abstract red dot, emphasizing once more that this phase concerns only the appearance of the dot itself.



In this phase, object recognition is entirely irrelevant, and no motion is yet extracted. The initial appearance of an environmental object—the emergence of the first deviant pixel—serves exclusively to delimit the vista. In the case of the Pong[3] game, we may know that the appearing dot must be the ball, but that is not the purpose of this phase. Its only function is to achieve the ultimate reduction of the visual field. This reduction does not arise from an initiated attentional shift but from an always-operational comparator system that is permanently active once the eyes are open.

When the eyes are open, this visual attention—just like the ever-present auditory attention—is always active, because our perceptual processes are implicitly developed to this goal. We are continuously capturing c.q. catching the entire vista, and so the conscious registration of this entering movement is direct. The transition to phase 1b therefore does not reflect the activation of a new attentional state but merely a refinement of this continuously operating baseline comparison process.

---

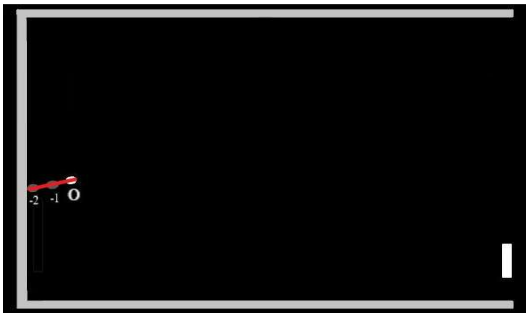
[2] https://www.researchgate.net/publication/397770908_The_Complete_Mechanistic_Clarification_of_The_Rubber_Hand_Illusion_-_Why_External_and_Internal_Hand_Perception_Are_Autonomous_and_Cannot_Be_Unified_Without_Motor_Action

[3] https://www.researchgate.net/publication/393795043_All_functional_perception_processes_in_relationship_to_PONG?_sg%5B0%5D=MWLOACaRyovQ-Qutu-udxE0V_rpgcFIWF1Y4-NxwJogcOAnYInwhV-4wfisP1HMKCFL5dxaUm33S-qNWu8uurkvq8fq5nxLwISak3zWw.qB--dcOSexBo7MiS-jTZMKzlmNsVZYKiQfBDJKPb8TNszhc7KxekMrUvJVSi67-CcMqJPv09D3uDPA-d5KAm8GA&_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6ImhvbWUiLCJwYWdlIjoicHJvZmlsZSIsInByZXZpb3VzUGFnZSI6InByb2ZpbGUiLCJwb3NpdGlvbiI6InBhZ2VDb250ZW50In19

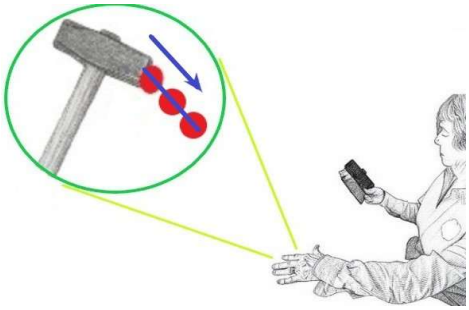
Although this phase involves a more detailed visual examination of the detected vista change, auditory perception remains continuously active as a general monitoring system. From an ecological perspective, anything in the environment can develop into a potential threat and must therefore be signaled as early as possible within the broadest possible field. The auditory system thus remains globally engaged, providing an ever-present 360° deviation detector that continues to operate while visual refinement progresses.

III. Visual Attention Phase 1b

Goal: Convert peripheral visual attention focused on a single point into peripheral visual attention focused on the *movement* of that point[4]. Visual attention phases 1a and 1b are tightly linked in temporal duration. Phase 1b, however, clearly *follows* phase 1a. Both phases are characterized by the fact that object recognition is not yet relevant.



1



2

Illustrations: *The core of the renewed paradigm demonstrated here is that every movement c.q. every perception of movement must originate from the previous positions of the action object. Which has never been sufficiently recognized in vision science. All perceptual registrations of the motion of the pong ball[5] or of the hammer are therefore always connected in line segment shapes, reflecting the fundamental conclusion that our perceptual system is primarily a comparison system for movement.*

In this subsequent phase, the system determines solely whether the detected point continues into actual motion and whether that motion may generate a potential point of intersection with the organism's own movement. At this stage the stimulus remains an UMO (unidentified moving object). The system compares only the earliest adjacent time frames[6].

From the manifest positions of the UMO, one can draw a line to the current position, enabling the organism to form a perceptual representation of the latent part of this *action trajectory shape*. If this perceptual projection suggests that the latent trajectory may intersect with the body's own movement, the organism will, out of caution, often withdraw the hand pre-emptively.

Although this phase involves a more detailed visual examination of the detected movement, auditory perception—together with the continuously operating detection of any vista-wide change—remains active as a general monitoring system, as mentioned within the previous paragraphs. From an ecological perspective, anything in the environment may develop into a potential threat and must therefore be registered as early as possible across the broadest possible field. The auditory system thus remains

[4] Visual attention, as used in this model, does not refer to a triggered resource or selective activation process. It denotes a continuously active baseline comparator present whenever the eyes are open.
[5] All functional perception processes in relationship to PONG – page 7
[6] The above shown images to illustrate this process probably show far too many positions of the pong ball and the hammer.

globally engaged, functioning as an ever-present 360° deviation detector that continues to operate while visual refinement progresses.

Motion Recognition versus Object Recognition

The explanatory model of the motoric movement action provides a definitive account of all attentional foci within visual perception. On the one hand, it confirms established findings in current vision science that object recognition is relatively slow, typically emerging around ~400 ms. On the other hand, it introduces several novel insights, most notably that the visual recognition of pure motion (>30 ms) is fast, continuously active, and always operational.

The model shows that motion recognition and object recognition form autonomous components of the attentional perceptual processes and do not exclude one another. Thus, even when object recognition becomes active, the detection of motion within the vista continues independently.

The primary aim of this article is to convey these autonomous principles. In this boxed section, it is explained that object recognition may already take place—or, at the very least, may be decisively guided—by prior recognition of pure motion. Before a hare or a duck has actually been identified as an object, such a conclusion may already have been inferred from the specific movement characteristics of the action object.

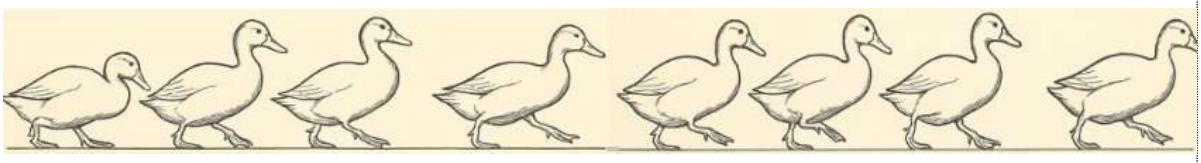
Illustrations[7]**:** *When it is established that a hare is moving within a vista, the fast and slow autonomous visual perceptual processes can be distinguished as well. The realization that something is moving c.q. that something occupies successive positions becomes apparent very rapidly, whereas the recognition of that moving "something" as an object, in this case a hare, emerges more slowly. These remain distinct and autonomous phases. However, through acquired cognitive knowledge, it becomes possible to associate a specific movement idiom with a particular action object. Thus, even if the hare has not yet been visually identified or recognized as such, one may already infer from its characteristic movement pattern that it must be a hare, or at least a hare-like entity.*

The perception of (external) motion is ecologically developed primarily to estimate, from the manifest positions P of an action object, the latent part of its action trajectory shape c.q. its near-future positions.. The action trajectory shape of an external environmental object may potentially yield a point of intersection with the observer's own egocentric action trajectory and must therefore be detected and evaluated.

At this stage, no account is given of how the recognition of motion c.q. the specific movement pattern of an action object may relate to object recognition and thereby allow object recognition to proceed more rapidly than explained above. The interaction between the (qualitative) perception of a

[7] Alfred Brehm: https://archive.org/details/brehmstierlebena11breh/page/n21/mode/2up

specific movement idiom and definitive object recognition c.q the perception of the specific environmental object is merely identified in this blocked section.



Illustrations: *When it is established that a duck is moving within the vista, the same two autonomous visual perceptual processes, a fast and a slow one, can be distinguished. As with the hare, most observers recognize the movement idiom of a duck almost immediately and will already have associated this idiom with a duck before the "abstract moving entity" is definitively recognized as such. It should be noted, however, that there is a substantial difference in tau-values between the rates at which a hare and a duck fill their respective action trajectory shapes, reflecting the pronounced difference in their movement dynamics.*

From these examples it follows that an interaction may arise which, under the same observation, can produce a clear difference in perceptual outcome. At the same time, the examples show that hybrid forms may occur, or that under identical observational conditions no interaction emerges at all between movement-pattern recognition and object recognition.

This latter conclusion can be reformulated into the ecological observation that early movement-idiom recognition can give rise to a powerful reduction process. On the basis of specific movement characteristics and object-specific *tau*-values, one can very rapidly conclude—negatively—that a moving entity is certainly not a hare when observing a duck, and conversely not a duck when observing a hare. This occurs "long before" object recognition, which typically takes place much later around ~400 ms and can therefore accelerate the object-recognition process by providing strong directional constraints.

This mechanism is directly comparable to the perception of zero motion. When entering a new environment, the entire vista is initially parsed into moving and non-moving points or pixels. We very rapidly recognize whether something is absolutely stationary and will not move at all. Although this does not yet reveal whether the perceived object is f.e. a coffee cup or a stuffed hare, it nevertheless produces a substantial reduction of possible interpretations.

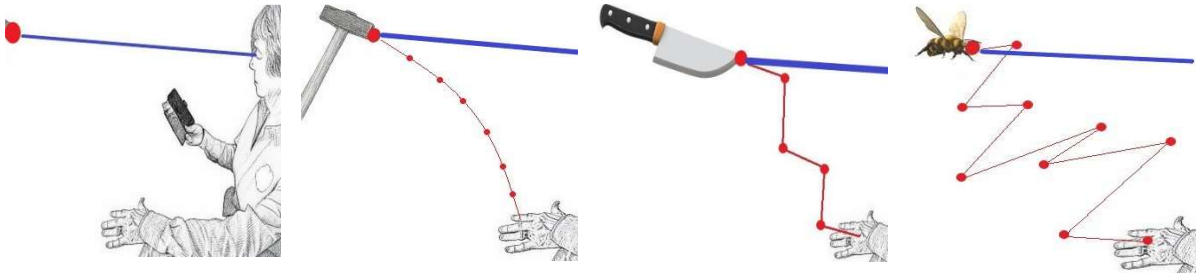
Illustrations: *The rapid observation of the movement idiom of an unfamiliar environmental object can contribute to a reduction process before a hare or a duck is actually visually recognized. In the same way, a substantial reduction can be initiated by observing absolute absence of motion. We recognize absolute zero motion in an as-yet unknown environmental object in such a way that we*

still do not know whether it is a coffee cup or a stuffed hare, yet a large number of possible interpretations are already excluded.

This reduction process operates independently of explicit object identification and conscious categorization and reflects an ecologically primary perceptual mechanism optimized for early constraint of possible action-relevant interpretations.

I. Visual Attention Phase 2

Goal: To transition from peripheral detection of confirmed movement (phases 1a and 1b) into direct focal vision, where the moving action object is recognized and the manifest portion of the action trajectory shape, including its *tau*-value, is constructed with greater precision..

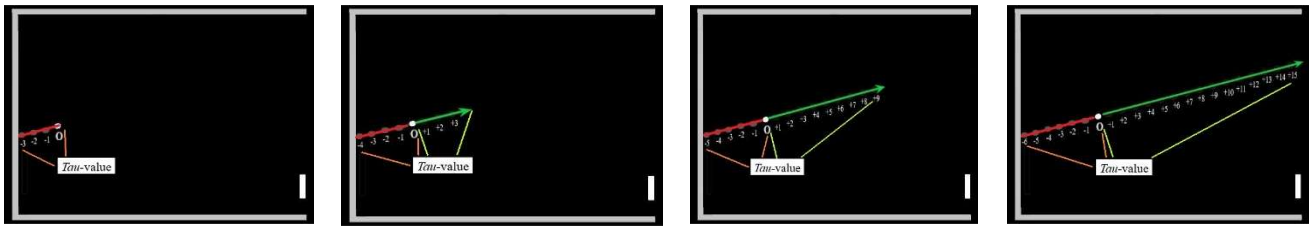


Illustrations: *The recognition of what exactly is moving provides additional information about both the shape and the tau-value of the relevant action trajectory. Thus, object recognition is essential, but it must be understood within an entirely different framework: its ultimate purpose remains to determine how the object's movement relates to the organism's (future) egocentric movement. So recognizing the future movement of a hammer, a knife, or a bee has far-reaching consequences.*

By explicitly identifying phases 1a and 1b as consciously distinct phases dedicated solely to the detection of movement, it becomes evident why object recognition emerges only later. Not every detected movement poses a threat, and therefore no energy needs to be expended unnecessarily—a principle that fits seamlessly within an ecological perspective.

However, when the movement detected in phases 1a and 1b exceeds a certain threshold, further visual investigation becomes necessary. Naming or identifying the specific action object must then provide essential information—*in relation to the developing action trajectory shape and its tau-value*—for determining how the trajectory will continue.

We possess extensive cognitive knowledge concerning the characteristic action trajectory shapes and *tau*-values of many different types of action objects. The explanatory model of the motoric movement action has demonstrated, across a wide variety of actions, how a perceptual representation of the latent part of the action trajectory—including its *tau*-value—arises directly from the manifest portion.



Illustrations: *The explanatory model of the motoric movement action shows that the movements of any conceivable action object (i.e., a pong ball[8], hammer, knife, or bee) must always originate from earlier positions. All perceptual registrations of actual movement (in red) therefore necessarily become part of a line-structured representation. Factually it follows that the latent portion of the trajectory (green) must arise from the manifest part. This applies both to its shape (x-axis form) and to its tau-value, related to time-to-contact (TTC).*

The greatest scientific breakthrough lies in the recognition that the successive positions of an action object must always originate from one another. As a direct consequence of the continuous visual attention to movement—i.e., the continuous comparison of adjacent time frames of the action object—all movement is necessarily perceived in the form of an action trajectory shape. It can therefore be established that the latent part of the trajectory must arise from the manifest part, both in its geometric form (x-axis structure) and in its *tau*-value, which is intrinsically linked to time-to-contact (TTC).

Consequently, we do not perceive the movement of the pong ball, hammer, knife or bee as a series of isolated points; rather, we perceive a moving object embedded within a single overarching projective structure. The perception of motion thus presupposes a unifying framework in which spatial succession is integrated into a coherent temporal experience, rather than being registered as a sequence of disconnected events.

Although this phase provides a more detailed specification of the detected movement, all other perceptual phases remain simultaneously operational. Newly arising sounds and any vista-wide changes indicating additional threats continue to be monitored without interruption, as the broader comparator architecture remains continuously engaged.

II. Visual Attention Phase 3

Goal: Continuously update the latent action trajectory, because predictive coding is only an approximation. The system must track the movement of the action object—and the required egocentric response—until the final moment of contact.

The perceptual processes described in the previous phase must remain active until the end of the action, because predictive coding of the latent part of the *action trajectory shape* provides only an estimate of the probability boundaries of future fluctuations. It is part of an optimization process that can

[8] All functional perception processes in relationship to PONG – page 8 & page 12.

predict the future positions *P* of the action object within a certain bandwidth with considerable accuracy, yet never exactly. Perception can make a reasonable assumption about the future, but never determine its precise unfolding; the action object will always deviate—infinitely—from the predicted segment.

Therefore, the incoming pong ball, hammer, knife, or bee must be visually mediated up to the very last moment, exactly as described in phase 2, and must be answered with an egocentric movement that either blocks the threat or avoids it (fight or flight). In the case of Pong[9], both the movement of the ball and the egocentric movement of the paddle must be mediated by the cortical streams until the final instant. With respect to visual attention, this means that the perceptual process does not cease until the entire attack or incoming movement is fully completed.

Although this phase is a continuation of the previous phase, all other perceptual phases remain simultaneously operational. Newly arising sounds and any vista-wide changes indicating additional threats continue to be monitored without interruption, as the broader comparator architecture remains continuously engaged.

[9] All functional perception processes in relationship to PONG.

References:

Amerio, P., Michel, M., Goerttler, S., Peters, M. A. K., & Cleeremans, A. (2024). *Unconscious perception of vernier offsets*. **Open Mind, 8**, 739–765. https://doi.org/10.1162/opmi_a_00145

Alais, D., & Burr, D. (2004). *The ventriloquist effect results from near-optimal bimodal integration*. **Current Biology, 14**(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Békésy, G. von. (1960). *Experiments in Hearing*. McGraw–Hill.

Carrasco, M. (2011). *Visual attention: The past 25 years*. **Vision Research, 51**(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Castanheira, J. da S., Shea, N., & Fleming, S. M. (2025). *How attention simplifies mental representations for planning*. **eLife** (advance online publication). https://doi.org/10.7554/eLife.108034.1

Cleeremans, A., Mudrik, L., & Seth, A. K. (2025). *Consciousness science: Where are we, where are we going, and what if we get there?* **Frontiers in Science, 3**, 1546279. https://doi.org/10.3389/fsci.2025.1546279

Cleeremans, A., & Tallon-Baudry, C. (2022). *Consciousness matters: Phenomenal experience has functional value*. **Neuroscience of Consciousness, 2022**(1), niac007. https://doi.org/10.1093/nc/niac007

Fleming, S. M., & Michel, M. (2025). *Sensory horizons and the functions of conscious vision*. **Behavioral and Brain Sciences**, 1–53. https://doi.org/10.1017/S0140525X25000068

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. Wiley.

Kahneman, D., & Treisman, A. (1980). *Changing views of attention and automaticity.* In R. S. Nickerson (Ed.), *Attention and Performance VIII* (pp. 29–61). Erlbaum.

Lee, D. N. (1976). *A theory of visual control of braking based on information about time-to-collision*. **Perception, 5**(4), 437–459. https://doi.org/10.1068/p050437

Luce, R. D. (1986). *Response Times: Their Role in Inferring Mental Organization*. Oxford University Press.

Miall, R. C., & Wolpert, D. M. (1996). *Forward models for physiological motor control*. **Neural Networks, 9**(8), 1265–1279. https://doi.org/10.1016/S0893-6080(96)00035-4

Michel, M. (2023). *How (not) to underestimate unconscious perception*. **Mind & Language, 38**(2), 413–430. https://doi.org/10.1111/mila.12406

Michotte, A. (1954). *La Perception de la Causalité*. Publications Universitaires de Louvain.

Posner, M. I. (1980). *Orienting of attention*. **Quarterly Journal of Experimental Psychology, 32**(1), 3–25. https://doi.org/10.1080/00335558008248231

Regan, D., & Gray, R. (2000). *Visually guided collision avoidance: Time-to-collision is faster than optical tau*. **Journal of Experimental Psychology: Human Perception and Performance, 26**(2), 303–321. https://doi.org/10.1016/S1364-6613(99)01442-4

Sanders, A. F. (1998). *Elements of Human Performance: Reaction Processes and Attention in Human Skill*. Erlbaum.

Shelton, J., & Kumar, G. P. (2010). *Comparison between auditory and visual simple reaction times*. **Neuroscience & Medicine, 1**(1), 30–32. https://doi.org/10.4236/nm.2010.11004

Simons, D. J., & Rensink, R. A. (2005). *Change blindness: Past, present, and future*. **Trends in Cognitive Sciences, 9**(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Stockart, F., Schreiber, M., Amerio, P., Carmel, D., Cleeremans, A., Mudrik, L. (2025). *Studying unconscious processing: Contention and consensus*. **Behavioral and Brain Sciences**. https://doi.org/10.1017/S0140525X25101489

Warren, W. H. (2006). *The dynamics of perception and action*. **Psychological Review, 113**(2), 358–389. https://doi.org/10.1037/0033-295X.113.2.358