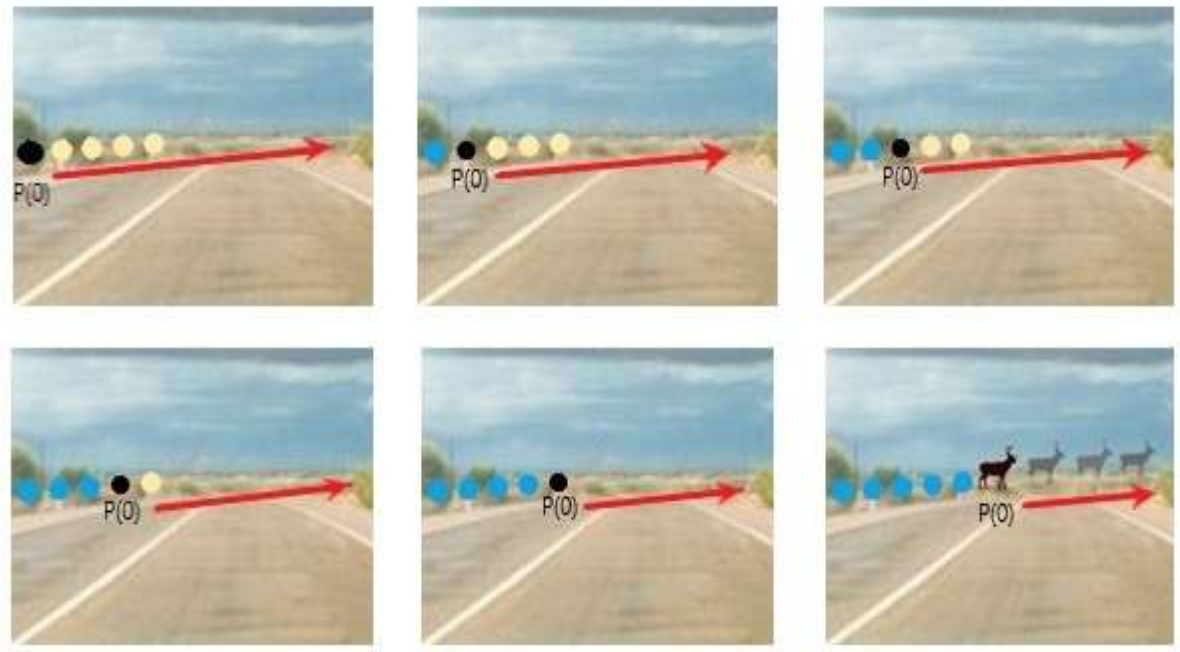
Beyond *Sensory Horizons and the Functions of Conscious Vision*: A Commentary on Fleming & Michel (2025)

Foundational Misconceptions, Theoretical Omissions, and Core Fallacies in Modern Vision Science



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

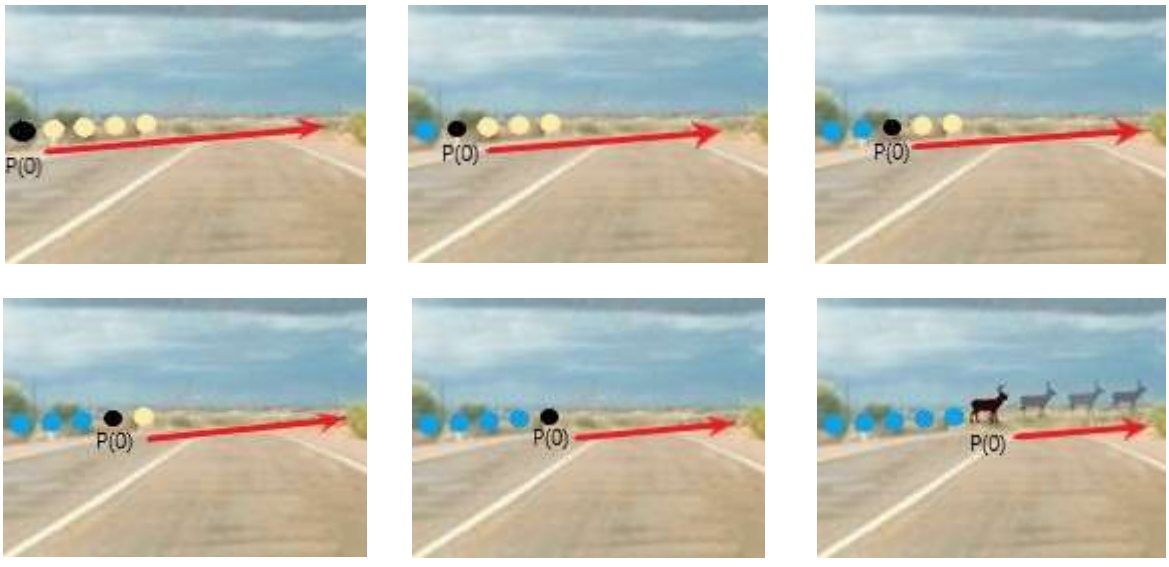
N.J. Mol
December 2025 ©

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

Commentary

Fleming & Michel offer an ambitious attempt to clarify the functions of conscious vision, yet their account remains incomplete precisely at the point where a mechanistic explanation is required. While we agree that conscious vision plays a central role in guiding adaptive behavior, the architecture they propose is functional rather than mechanistic—reflective rather than generative. What is missing is the underlying process that produces conscious visual content in the first place.

The target article repeatedly characterizes conscious vision in terms of functions—horizon expansion, integration, flexibility—but never identifies the mechanism that generates conscious experience itself. Without such a mechanism, the theory cannot distinguish necessary components from derivative phenomena. A viable account must explain *how* visual content becomes conscious, not merely *what* consciousness enables organisms to do.



Illustrations: *It is a factual constraint that the movement of any conceivable action object must arise from its own successive positions—that is, its positions are always embedded within a single continuous trajectory. The explanatory model of the motoric movement action demonstrates that our perceptual systems are primarily attuned to this movement and remain continuously active. Even when the*

action object has not yet been recognized, its movement is already perceived. In the case of the moose example discussed by Fleming & Michel, the organism perceives movement solely by comparing the initial positions of the yet-unidentified object, thereby forming a perceptual representation of the latent portion of the action trajectory, including both its geometric form and its tau-value. The subsequent recognition of the moose becomes merely one aspect within the overall perceptual process and remains entirely subordinate to refining the movement estimate within the perceptual representation of the latent action-trajectory shape.

A movement-comparison architecture provides this missing foundation. Conscious visual content arises directly from the continuous comparison of adjacent temporal frames within the perceptual system. This comparator, permanently active as long as sensory organs are engaged, generates the very differences that appear as conscious experience. Consciousness is therefore not an additional layer added to perception; it is the experiential form of perceptual comparison itself.

Consequently, Fleming & Michel's distinction between conscious and unconscious vision collapses under a comparator architecture. Once the comparator is active, all detected differences are consciously available by design; there is no "unconscious vision" that later becomes conscious. Their model presupposes intermediate layers of processing that ecologically do not exist.

The authors also treat conscious vision as a functionally useful capacity rather than as an ecologically grounded survival necessity. Evolution does not select for abstract horizon expansion—it selects for the earliest possible detection of environmental change. Movement, not object identity, is the universal referential dimension of the perceptual system. All later visual capacities emerge from progressively refined stages of movement detection, from coarse global deviation to precise trajectory estimation. This foundational principle is absent in their account.

By grounding consciousness in an always-active temporal comparator, perception becomes inherently conscious, movement-first, ecologically optimized, and mechanistically defined. This framework specifies the architecture that Fleming & Michel gesture toward but do not articulate: conscious vision is not a functional add-on, but the direct experiential manifestation of an underlying movement-comparison mechanism.

Abstract

Current vision science still lacks a coherent mechanistic framework, largely because it is built on longstanding misconceptions and categorical errors regarding visual attention. Fleming & Michel's functional account exemplifies this problem: it offers descriptions of what conscious vision enables, but no explanation of how conscious visual content is generated. This article presents the missing architecture: an ecologically grounded mechanism in which auditory and visual perception originate from a single primordial survival function—the continuous detection and evaluation of environmental change c.q. movement, early enough to act.

Perception is therefore not capacity-limited or intermittently "activated," but a permanently operating comparator that evaluates differences across adjacent temporal frames. Consciousness is the direct experiential outcome of this comparator activity: every detected change across frames is consciously perceived at the moment it occurs, making consciousness an inherent property of the perceptual mechanism rather than an added cognitive layer.

Within this architecture, the organism responds to approaching objects through 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.

Although sequentially described, all phases remain continuously active as parallel comparator streams. This multilayered structure reframes consciousness itself as the phenomenal expression of an always-on ecological comparison system fundamentally dedicated to detecting, evaluating, and responding to movement—precisely the mechanistic foundation absent in Fleming & Michel's proposal.

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

Current vision science still lacks a coherent mechanistic framework, largely because it is built on longstanding misconceptions and categorical errors regarding visual attention. These shortcomings reflect not only conceptual confusion but also significant theoretical omissions and a persistent lack of insight into how perceptual systems actually operate. This article presents the first comprehensive account that completely resolves these issues. The explanatory model of the motoric movement action (2016) provides a unified and ecologically grounded structure for all visual functions and offers definitive evidence for their underlying mechanism. It demonstrates that both auditory and visual perception arise from a single primordial evolutionary imperative: the need to detect environmental *change* c.q. *movement* early enough to survive. This function is not ancillary but constitutes the organism's most fundamental and continuously operating survival system.

As a result, the perceptual system requires a mechanism capable of functioning continuously, rapidly, and with minimal computational cost. This requirement is fulfilled by the ongoing comparison of perceptual images across adjacent temporal frames. What begins as a permanently active 360° auditory deviation detector, the organism's first line of survival, expands into an intermediate vista-based peripheral field and finally into high-resolution focal vision. Each stage refines movement detection from coarse global deviation to highly precise trajectory estimation. Subsequent visual capacities are therefore not independent modules but elaborations of a single, original motion-comparison mechanism that forms the foundational architecture of the entire perceptual system.

Within this architecture, the organism's response to an approaching object unfolds through five distinct phases: one auditory phase providing a global deviation detector, followed by four visual phases of progressively increasing specificity. Visual phase 1 registers the initial deviation—the first divergent pixel anywhere in the vista. Phase 2 determines whether this deviation reflects genuine motion and whether its latent continuation may intersect with the organism. Phase 3 transitions to focal vision, where the moving object is identified and the manifest part of its action trajectory shape and *tau*-value is used to project its latent continuation. Phase 4 refines this projection through predictive estimation of future fluctuation boundaries, enabling time-critical egocentric movement preparation until the approaching threat is neutralized.

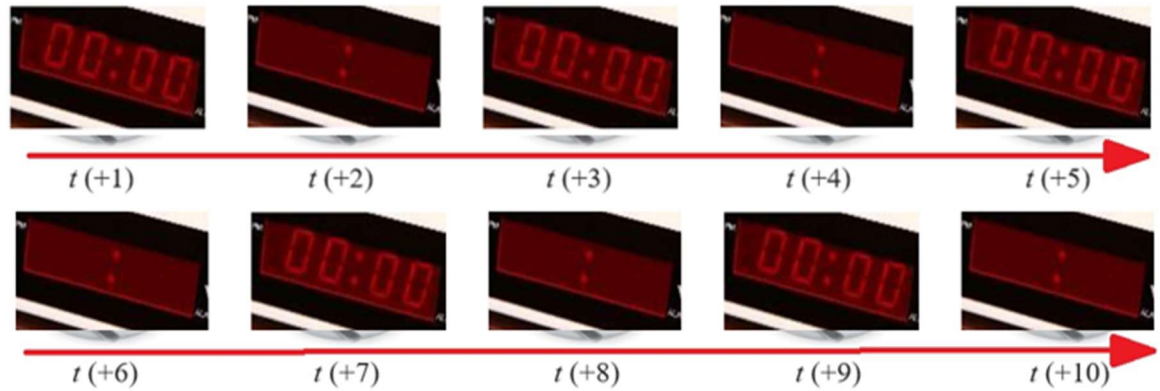
Although described sequentially, these stages do not replace one another. Once active, each phase remains continuously present as an autonomous comparator stream, maintaining its specific informational contribution while later phases add further refinement. Recognizing this architecture resolves

longstanding paradoxes in the vision literature and reveals visual attention as an ecologically optimized, multilayered system fundamentally dedicated to detecting, evaluating, and responding to movement.

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.

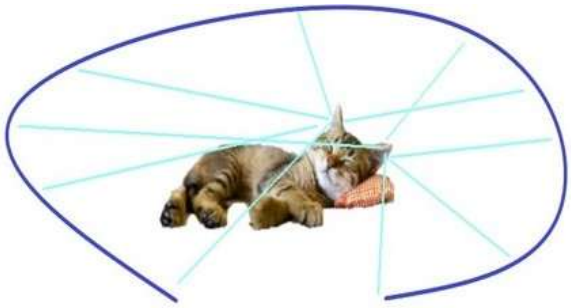


This provides definitive scientific evidence that our perceptual organs/systems are, above all, comparison systems. Such comparison systems operate continuously; they do not require activation and therefore impose no reaction-time limitations on their earliest perceptual stages. Even more crucial is 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.

I. Auditory Phase



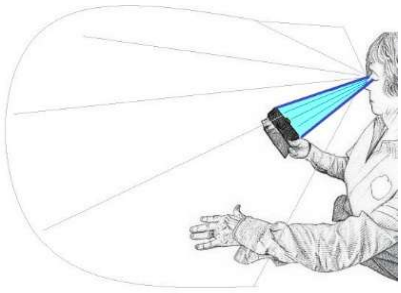
1

2

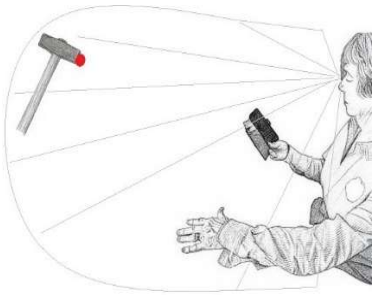
The primary level of motion detection encompasses the auditory attentional processes. They cover a 360° auditory field, which marks the lateral positioning of the ears—situated at the sides of the head, in relation to the eyes, close to the brain, and always active (even during sleep). During sleep (fig. 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

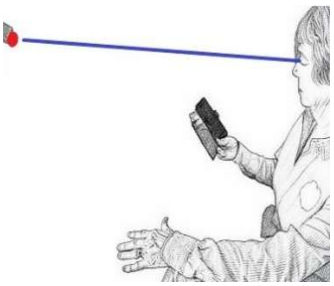
Goal: Reduce peripheral visual attention from the entire vista to one point c.q. an extremely restricted region[1].



1



2

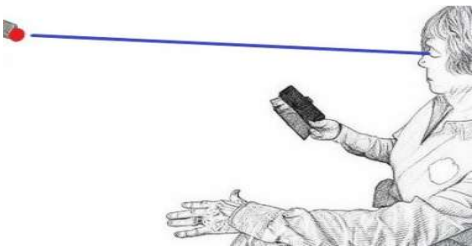


3

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.

[1] The ecological primary objective within an optimization 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



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.

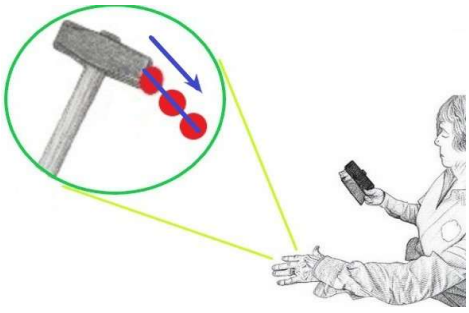
When the eyes are open, this visual attention—just like the ever-present auditory attention—is always active, because our perceptual processes are implicitly developed to this goal. We are continuously capturing c.q. catching the entire vista, and so the conscious registration of this entering movement 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.

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

[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

[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.

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.*

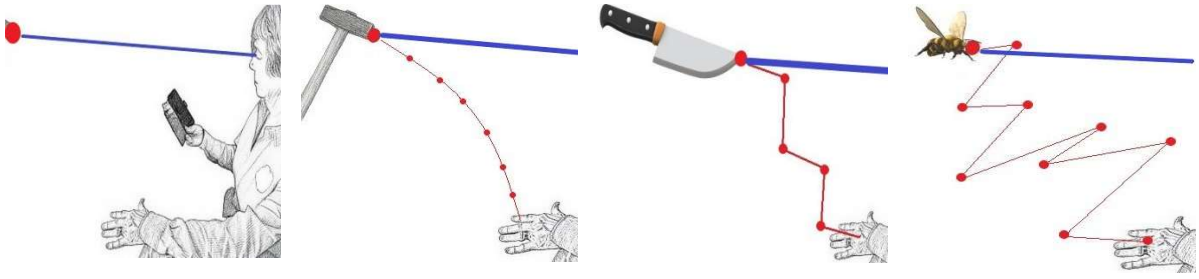
In this subsequent phase, the system determines solely whether the detected point continues into actual motion and whether that motion may generate a potential point of intersection with the organism's own movement. At this stage the stimulus remains an UMO (unidentified moving object). The system compares 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 globally engaged, functioning as an ever-present 360° deviation detector that continues to operate while visual refinement progresses.

IV. 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.*

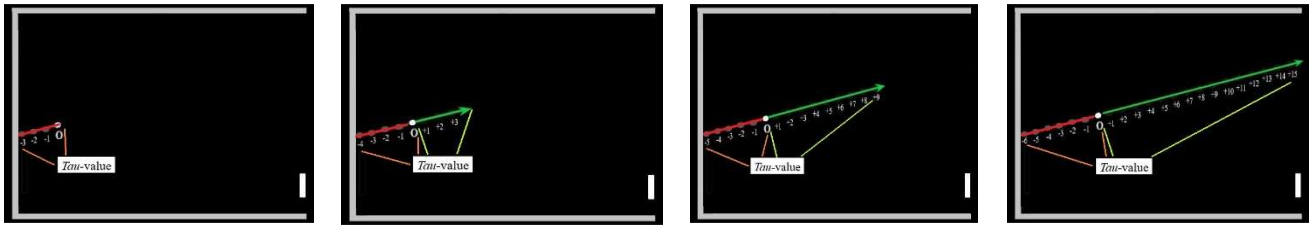
By explicitly identifying phases 1a and 1b as consciously distinct phases dedicated solely to the detection of movement, it becomes evident why object recognition emerges only later. Not every detected

[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.

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[7], 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.

V. 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.

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

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 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[8], 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.

[8] 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