Tattva – Journal of Philosophy
2026, Vol. 18, No. 1, *107-127*
ISSN 0975-332X/https://doi.org/10.12726/tjp.35.7

Enactive Perception vs Zen Perception: A thought experiment using Muller-Lyer Illusion

Navneet Chopra*

Abstract

Enactive Cognition characterises perception as an 'emergent' phenomenon - not something like a faithful passive representation of pre-existing objective reality, but as something emerging from the coupling between sensory and motor experiences, from a pragmatic need to act in the world under the (action) affordances offered by the environment to the organism, and from dynamically coupled interactions between the person and others under the normativity of the situation one is in. All this suggests that no objective reality or objective perception is possible. On the other hand, Zen prescribes abandoning the encumbrances like competition, greed, fear, jealousy, etc., to perceive reality objectively 'as it is'. This stance means endorsing realism, i.e., believing that a real reality exists, and a 'true' perception is possible. Thus, the two perspectives stand opposite to each other. I reflect on this problematic situation through a thought experiment using the Müller-Lyer illusion, and raise a question - "Will a Zen master perceive the Müller-Lyer illusion?" And I think they will perceive the illusion because the perception of the illusion is the mark of being an 'enactive being', well acclimatised with one's environment, from which even a Zen master can't escape, if they are to achieve a judicious sense-making of the world.

Keywords: Enactive cognition, Fodor, affordances, emegence, realism, Zen, Müller-Lyer illusion

* Department of Philosophy, Delhi University, New Delhi, India; nchopra@philosophy.du.ac.in

Introduction

Who are we? Cognitive agents driven by some complex abstract rule (likean AI system), or Heideggerian Daseins? We are certainly not the former, and Heideggerian philosophy is concerned with human existence, and it characterises human existence with the term "Dasein", which means "being (out) there". If we are Dasein, then for the perception of the world, or the perceptual *meaning* in the world, the world "matters" to us! How do we see things or perceive the world, i.e. what is the underlying cognitive mechanism for this cognitive capacity available to us?

Fodor (1983) and Pylyshyn (1984, 1999) proposed certain visual modules to facilitate human perception, which are supposed to be "cognitively impenetrable," which means their output is unaffected by the 'beliefs' of the person about the state of the world. Fodor (1983) articulated a strong version of the modularity thesis, claiming that certain cognitive systems, especially the visual system, operate in a modular fashion. According to his model, these systems are domain-specific, operate quickly and automatically, and are informationally encapsulated, i.e. they do not access or integrate information from other cognitive domains such as belief systems or background knowledge, therefore they are "cognitively impenetrable", i.e., perceptual hypotheses can only be confirmed by information within the perceptual system itself, not by what the organism knows more broadly (Fodor, 1983, pp. 69–70; Fodor, 1985, pp. 1–5).

Fodor's central claim, in a way, is epistemological. It claims, in essence, that if perception were thoroughly pervaded by cognitive states, then its objectivity would be threatened, making perceptual knowledge indistinguishable from subjective elements like biases, prejudices, expectations and anticipations. Building on Fodor's foundation, Pylyshyn (1999, 2003) focused on early-vision as the exemplar of such modular processing. He argued that early visual processing is cognitively impenetrable. Aligned with their views, Raftopoulos (2014) claims that nonconceptual content (NCC) in the "early-vision" is cognitively impenetrated (CI) or encapsulated by the higher cognition (or affective elements).

But what exactly is the early vision? Broadly speaking, early vision involves feature detection, segmentation, and surface organisation during the first ~100 milliseconds, occurring in the early layers of the visual system, primarily in the retina and the primary visual cortex, V1. These early stages include low-level processes such as edge detection, colour processing, motion detection, and other basic visual feature extraction. Early vision is thus characterised by bottom-up cognitive processing, where visual data or information is received from the environment and processed in a relatively automatic and pre-attentive manner. It is considered to be more sensory-

driven and less influenced by high-level cognitive processes, like goal or purpose, attention, memory, anticipation, and expectation, etc. This claim, however, is highly contested by various studies, which show that even early vision is influenced by top-down higher cognitive and affective factors, as will be demonstrated in this paper.

According to Pylyshyn, early vision delivers a belief-independent representation of the visual world, which forms the foundation of visual perception prior to the influence of cognition or attention. He says: **"The early vision system is encapsulated from cognition, or to use the terms we prefer, it is cognitively impenetrable."** (Pylyshyn, 1999, p. **344**). This early-stage impenetrability, he claimed, is sufficient to preserve a modular structure within the visual system, even if later stages of visual experience are subject to cognitive influences. Like Pylyshyn, Raftopoulos (2001, 2009) agrees that early vision is cognitively impenetrable, while late vision is influenced by top-down factors like attention and concepts. However, our concern needs to be with the *overall* cognitive process and its final outcome - the final perceptual content or object one sees, or the final perceptual experience of the object one sees. It simply doesn't matter if a small early part of this process is shown to be impenetrable. One makes decisions or acts in accordance with what one achieves as the final perceptual object/perceptual experience. For example, an electrician would not choose to use a shorter ladder to reach a bulb on a distant electric pole, mistakenly perceiving it as lower than it actually is. Similarly, one does not worry that a friend has shrunk when they appear smaller from a significant distance. The partial insulation proposed by Pylyshyn and Raftopoulos is inadequate because the content of perception that actually guides action and belief is not fixed at an early stage. Their line of defence rests on a problematic assumption that establishing cognitive impenetrability in the early stages of vision suffices to defend the modularity of visual perception as a whole. This is misleading because perception is a temporally extended, hierarchically structured process in which later stages significantly shape conscious experience. Human judgments, decisions, and behaviours rely on the final perceptual output, which integrates both bottom-up and top-down influences. If cognition can influence this final stage—as empirical studies increasingly suggest—then the epistemic and functional autonomy of perception collapses. Lupyan (2015) argues that predictive coding models of perception entail top-down effects at all levels of processing, thereby refuting the idea of an impenetrable, sealed early vision. Most notably, Albert Newen and his associates challenge the modular view by emphasising the *situatedness* and *conceptual enrichment* of perception. They argue for a constructivist model of perception in which prior knowledge, social context, and conceptual background are integrated into perceptual processing even at

relatively early stages. Rather, Newen & Vetter (2017) claim that even early visual processing operates under top-down processing, under which a spectrum of cognitive influences across different stages of perceptual processing, from very early to late stages, and from low-level feature detection to high-level concept-based interpretation, needs to be accepted. They document various empirical studies demonstrating extensive *structural feedback* in the brain and *top-down influences* on early visual processing. They catalogue specifically varieties of cognitive penetration—ranging from subtle adjustments in feature discrimination to profound changes in object recognition all driven by beliefs, expectations, emotional states, and conceptual frameworks, and contend that empirical evidence for feedback loops and predictive processing undermines the plausibility of an encapsulated or sealed "early vision" module and suggests instead that cognition continually interacts with perception in a dynamically structured, content-rich way.

Further, Newen and Schlicht (2009, pp. 213, 236)[1] criticise that even in the field of social perception, such as face recognition and emotion perception, which are instances of relatively early perceptual processes, are also influenced by conceptual knowledge, background beliefs, and emotional states. They argue that understanding other minds involves multiple strategies, including non-conceptual direct perception and conceptual person-models, depending on context and familiarity where conceptual knowledge and expectations can influence emotion perception, challenging the solely bottom-up, encapsulated mode of vision in the field of social perception, and aligning with the view that top-down processes play a constitutive role in shaping what is perceived, e.g., facial emotion judgments under conceptual priming effects. This makes the case for rejecting the idea of a purely bottom-up, encapsulated visual system and proposing instead a *constructivist* view of perception, where top-down processes play a constitutive role in shaping what is perceived, particularly in social perception and, more generally, in perception.

Thus, early-vision modularity, as proposed and defended by Pylyshyn and Raftopoulos, fails to capture the holistic and interactive nature of perception. Perceptual processing is shaped by top-down cognitive influences, as emphasised by Newen and other scholars, showing that perception is deeply integrated with cognition. As a result, even if the early stages of the cognitive process responsible for perception show some modular features, the final, conscious content of perception is cognitively penetrable and not modular. While Fodor accepts the contextual influences in belief formation in inferential mode, but not in observation mode, these studies show that the influence of top-down processes exists even in the observation mode, unlike what Fodor believed. Rather, asking whether cognition penetrates "early" or "late" perception is misguided, since perception does not

decompose into epistemically autonomous stages in the first place. What matters is the overall structure of the perceptual system and the nature of the perceptual content it delivers. If perceptual experience systematically varies as a function of the subject's beliefs, goals, emotional states, or conceptual capacities, then the impenetrability thesis fails, regardless of whether some isolated subroutines remain encapsulated.

Enactivist Perspective

On the other hand, Heideggerian phenomenology, in view of Taylor (2007), informs us about certain *constitutive* Background Conditions – embodiment (i.e., bodily existence), culturality, historicity, instrumental or pragmatic relation with the objects, which roughly corresponds with (Wittgensteinian) notion of Form of Life. Against this backdrop, this paper proposes the existence of not mere "penetration", but "constitution" of perception under these background conditions. Perception is not just a modular activity, which is "encapsulated" as Fodor proposes, and has been desperately tried to be defended by his allies Pylyshyn, and later by Raftopoulos. This reflects rather a naïve "analytic" way of looking at how perception happens for us, the Daseins, who are not robots driven by an AI. Hence, this paper registers a protest against the *artificial* modularity approach in cognitive science.

From an enactive or interactionist perspective, perception emerges as a dynamic process involving the interaction of a living organism with its environment. Rather than passively receiving sensory input, the organism actively engages with its surroundings, learning from previous experiences and integrating changes in sensory information with movements toward or around objects. This perspective is supported by experiments such as those conducted by Bach-y-Rita, where tactile-visual substitution devices allowed blind individuals to "see" by translating visual information into tactile stimuli, demonstrating how perception is shaped by active exploration (Bach-y-Rita, 1972). Similarly, the classic kitten carousel experiment by Held and Hein (1963) showed that visual perception develops through active movement; kittens allowed to move freely developed normal vision, while those passively carried did not. These findings illustrate that perception involves anticipation and expectation, as the organism actively interprets and makes sense of incoming sensory stimuli, rather than merely representing an objective, external reality.

Further, the active and interactive enactive mind completes the *impoverished* data with the nearest meaningful gestalt. Here are a few examples. First, consider an ambiguous word – C◒T, where the lower portion of the letter is occluded, making it unclear whether it is O or Q, and so it could be either COT or CQT. But it is perceived as COT since that is the only meaningful gestalt or meaning word we have in standard English. Second, a

person can perceive a tourist tent as a wild bear when one already expects and anticipates it while crossing through a forest at twilight (Nolen-Hoeksema et al., 2009, p. 176). This paper dubs it as "tent as bear" phenomenon. Similarly, the perception of Muller-Lyer illusion can be considered as an example of such an interactive process. Fodor's theory of modularity suggests that certain cognitive processes, like visual perception, are "encapsulated" and operate independently of other mental processes. This means they function without influence from higher-level cognitive functions such as beliefs, desires, or knowledge. These visual modules, thus, are characterised by domain specificity, speed, and informational encapsulation, according to Fodor (1983). He uses the Müller-Lyer illusion to support his argument for modularity, highlighting that even when people know two lines are of the same length, they still perceive one as longer. For Fodor, this indicates that visual perception is modular, as it produces the same outcome regardless of contradictory higher-level knowledge.

However, the susceptibility to the Müller-Lyer illusion varies across different cultural contexts, challenging the idea of encapsulation. Studies by Segall, Campbell, and Herskovits (1966) show that individuals from non-Western cultures, particularly those not accustomed to "carpentered" environments—where right angles and straight lines are prevalent—are less likely to be fooled by this illusion. The Zulu and Navajo tribes, who live in environments with rounder structures than rectangular ones, exhibit a reduced susceptibility to the illusion (Pedersen & Wheeler, 1983, pp. 3-6). This suggests that visual perception involves an interaction between perceptual input and (physical) cultural experiences, pointing to a more interactive model of perception incorporating top-down and bottom-up processes. Further research by Henrich et al. (2010) underscores that perception varies across different cultural groups, challenging the universality assumed by modular theories. The failure of the Müller-Lyer illusion in diverse cultural contexts undermines Fodor's claim for an encapsulated system of perception and supports a dynamic, feedforward-feedback model of perception.

Further, Fodor seems to mischaracterise cultural and contextual factors as mere "theories" by suggesting that perception operates in a theory-neutral, encapsulated manner, where the mind passively constitutes perception without external influences. However, this view overlooks the reality that cultural and contextual factors are not just theoretical constructs but foundational "background conditions" that shape how we perceive and interact with the world. From a Heideggerian standpoint, perception is understood as part of *being-in-the-world*; this means that the background conditions, such as cultural practices, bodily action, and instrumental-pragmatic relationships with objects, fundamentally shape how we engage with and make

sense of the world (Heidegger, 1962; Taylor, 2007). These elements play a pragmatic role in our everyday activities, influencing how we perceive and interact with objects. Similarly, Wittgenstein's concept of "forms of life" emphasises that our shared practices and cultural backgrounds are integral to meaning-making, grounding our understanding of a word or phrase within a specific cultural-historical context (Wittgenstein, 1953). Therefore, cultural factors should be seen as integral to perception, not merely theoretical overlays, challenging Fodor's claim of a theory-neutral, modular perception.

J.J. Gibson's (1979) "direct perception" approach also contradicts conventional cognitive theories that view perception as a brain-based thinking activity. According to Gibson, perception is the outcome of interaction between the organism and its environment, where information is directly picked up from the surroundings without the need for internal mental representations. The idea of "affordances"—the action possibilities that objects provide to an organism according to their physical capabilities is actively employed here. For instance, a handle allows grabbing, a ladder allows climbing, a chair allows sitting, and so on. These affordances result from the interaction between the item and the perceiver's capacity for action rather than being solely attributes of the objects.

Gibson's emphasis on the direct, action-oriented nature of perception aligns closely with Heidegger's concept of "ready-to-hand" (*Zuhandenheit*) from *Being and Time* (Heidegger, 1962). For Heidegger, objects in the world are mainly regarded as tools or equipment that are characterised by their usefulness and relevance in daily actions rather than as inert entities with intrinsic features. When the hammer is utilised for pounding, it is viewed as "ready-to-hand" rather than as an abstract item. Gibson's affordances also imply that perception and action are intrinsically linked, with objects being perceived according to the action-oriented pragmatic value to the perceiver. Together, these views make the case that perception is inherently tied to the actions possible with the objects perceived for the perceiver.

In contrast to the conventional modular, representational theories of perception, Gibson and Heidegger both stress that perception is based on an organism's active engagement with its environment, modified by practical activities, objectives, and embodied capacities (Gibson, 1979; Heidegger, 1962). Human perception is strongly influenced by the desire for effective action, indicating that perception is an active, goal-directed process rather than just the passive receiving of sensory data. This concept is consistent with the enactive approach, which holds that perception is intrinsically linked to an organism's capacity for successful navigation and interpretation of its surroundings (Noë, 2004). The functional part of

perception is highlighted by the fact that successful action, like reaching, gripping, or moving, depends on perceiving objects in a way that permits these activities.

The "dwarf couple" phenomenon, a perceptual illusion in which the same pair appears differently depending on their position in a visual context, is a conspicuous example of this. Even though they are objectively the same height, the couple appears dwarf-like when viewed from the front of the hall, but they appear normal when viewed from the back. The visual system's perception of perspective and depth cues causes this illusion. In addition to representing the "objective" dimensions, the brain interprets visual data to produce a perception that conforms to our preconceived notions about how people and things should seem in relation to their environment. This illustration shows how the brain's expectations based on prior experiences and the surrounding environment interact dynamically with sensory inputs during perception. The visual system increases the couple's apparent size based on its interpretation of their height in relation to their spatial position, demonstrating that perception is more than simply passively taking in images; it also involves creating a meaningful and action-relevant interpretation of the world around us. The notion that perception is essentially designed to promote effective action is supported by the brain's capacity to modify and misinterpret visual information in response to environmental circumstances. The point being made here is illustrated in Figure 1, shown below.

Figure 1: "Dwarf" Couple (circled and shown by a downward arrow)

Explanation of the Case of Muller-Lyer Illusion:

The following figure 2 shows the Muller-Lyer illusion, where the upper and lower lines trapped between two kinds of arrowheads, one pointing outwards and the other inwards, appear to be different in length when actually they are of the same length! What explains this illusion?

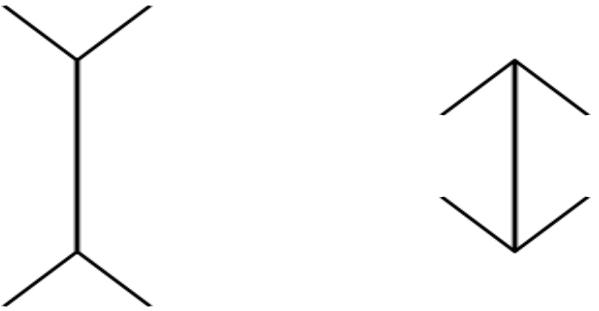

Figure 2: Muller-Lyer illusion

Contrary to the Fodorian explanation of the appearance of the Muller-Lyer illusion as a result of 'cognitively impenetrable' visual module, the illusion can be read as an example of the interaction of general cognition (which takes account of the surrounding environment as a context) with the visual stimuli corresponding to the figure under observation. This *interactive* explanation uses two phenomena for its explanation – the phenomenon of 'size constancy' and 'living in *rectangular* carpentered houses', which creates an environment for the object under observation (i.e. two lines).

Jean Piaget (1954, pp. 11–13) proposed the idea of "size constancy," which is the result of an infant's active interactions with their surroundings as they learn about the world. The knowledge that things retain their true size even as their distance from the observer varies is known as size constancy. When an object is farther away, for example, it may appear smaller, but the infants gradually learn through their interactions with the objects, say a ball, that this apparent diminution does not correspond to a change in the object's actual dimensions. This developing perceptual ability, known as size constancy, might be crucially involved, one can argue, for navigating the world effectively. It enables a person to judge the actual size of objects regardless of their distance from the observer. For example, an ordinary adult does not perceive a fifteen-foot-tall electric pole as being only two inches tall simply because it is viewed from a great distance. The

perceiver *mentally elongates* the visibly diminished size of the pole to compensate for its loss of apparent height under their learning of *size constancy* in infancy through their interactions with the world.

The *rectangular* carpentered houses, characterised by rectangular cross-sections, contain four walls and four edges inside, as well as four external edges. When observing an inner vertical edge of such a structure, the corners at the top and bottom represent the joints with the roof and the floor, respectively. This architectural arrangement contextualises the perception of the inner and outer edges of the room, inward-pointing arrowheads for the inner edge and outward-pointing ones for the outer edge, as can be seen in Figure 3 below.

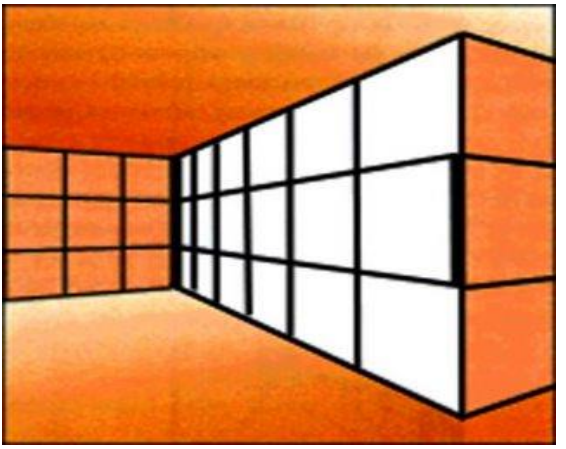

Figure 3: Building corners cueing the visual perception

The alternative (to Fodorian) explanation for the Müller-Lyer illusion posits that it arises from an interaction between two phenomena: our mental elongation of visual images of objects seen from a distance, which serves to maintain size constancy, and our adaptation to living in rectangular homes—specifically, those with rectangular cross-sectioned rooms. This interaction is critical in how we perceive the orientation of the arrowheads, contextualising the two lines in the illusion.

When viewing the line with inward-directed arrowheads, it is interpreted as the inside corner of a rectangular room that extends away from us. Conversely, the line with outward-directed arrowheads is perceived as representing the relatively close corner of a rectangular object, such as the corner of a building. This cognitive processing is influenced by our prior experiences with distance cues; we tend to assume that the outside corner (associated with the outward arrowheads) is closer than the inside corner (associated with the inward arrowheads). As a result, we deduce that the

inside corner must be longer, leading to the perceptual distortion central to the Müller-Lyer illusion.

Empirical evidence supporting this explanation can be observed in studies of the Zulu tribe, who are raised in circular homes characterised by cylindrical walls. According to research, people from this tribe are much less prone to the Müller-Lyer illusion than those from cultures where most buildings are constructed with rectangles and right angles. This disparity supports the notion that our experiences with architectural forms affect how we interpret visual stimuli and demonstrates how the contextual context impacts perceptual processes (Segall et al. 1966). Our adaptations to particular surroundings can significantly change our perceptual experiences, according to the findings, which also emphasise the importance of cultural variables in perception. It might be tempting to assume that exposure to culturally specific visual stimuli is necessary for "correct" environmental perception. However, from an enactive standpoint, perceptual systems are attuned to the regularities of their physical environment, and the persistence of certain visual illusions reflects this ecological adaptation rather than *familiarity* with particular stimuli. The relevant environment for perceptual attunement is the physical and ecological context, not social-cultural conventions. The crucial point is that the inference does not depend on familiarity with particular stimuli, but on the fact that perception is enacted by an agent who is already appropriately attuned to, and able to act successfully within, their environment. From an enactive standpoint, the persistence of certain visual illusions is precisely a mark of perceptual systems that are well-adapted to their environment.

In Heideggerian philosophy, the observations regarding perception and meaning make coherent sense, unlike in Fodor's perspective, which is rooted in a limited and somewhat distorted metaphysical worldview. Fodor's view treats the world as a collection of discrete entities or objects that derive their meaning from external, abstract rules rather than from inherent qualities of the objects themselves. In contrast, Heidegger posits that objects are not isolated entities; rather, they are interconnected, and their meanings are shaped by pragmatic needs. According to Heidegger, the "ready-to-hand" concept indicates that objects are imbued with meaning through our engagement with them from the outset, particularly when we "care" for or utilise them in practical ways. This inherent pragmatic connection provides the causal or constitutive role in our perception of these objects, infusing them with meaning based on their potential for action (Heidegger, 1962).

Gibson (1979) goes even one step further and ascribes the action possibilities or affordances as part of their meaning even when one merely sees

or perceives them. When a hammer is perceived lying on a table rather than held in hand, it is recognised as a tool for potential use, such as driving nails into walls or breaking objects. In contemporary discussions, we can assert that this enactive need is present from the very first encounter with an object like a hammer. This phenomenon extends beyond man-made tools; it applies to any object, whether natural or artificial, that can be utilised in some manner. The potential for use is inherently part of the object's meaning, rather than being dictated by external rules.

The groundbreaking work of Bach-y-Rita et al. (1969) with congenitally blind individuals further emphasises the dynamic nature of perception. In his studies, when cameras were mounted on the heads of blind participants, their ability to navigate and interpret visual stimuli illustrated how their prior tactile and auditory experiences influenced their subsequent understanding of visual information. These findings suggest that vision is not simply a passive reception of data but is actively shaped by an individual's prior experiences and interactions with their environment. Nonetheless, the phrase "direct perception," as employed by certain academics such as Hutto (2005), can be problematic and ambiguous. When analysed from a subpersonal perspective—the area of cognitive science—it may be viewed as problematic, even though it may accurately characterise the experience of perception at a personal level. Perception is not a simple or instantaneous phenomenon at this subpersonal level; rather, it is the result of a complex interaction of cognitive processes that are outside our direct access and that underlie what is phenomenologically referred to as "direct" observation.

In Barsalou's embodied account of cognition, meaning is not solely encoded within the sensory cortex but also involves the motor cortex and affective regions of the brain (Barsalou, 2008). This suggests that our understanding of objects, such as a hammer, is deeply intertwined with our potential actions and emotional responses to them. According to this viewpoint, perception is a dynamic process that involves both our physical interactions and the environment in which we encounter objects, and it is not only the outcome of sensory information.

Furthermore, Fodor's belief that concepts are only instruments for thinking and not for doing has been criticised by academics such as Prinz and Clark. They contend that our cognitive processes are essentially action-oriented and that concepts are crucial in directing behaviour (Prinz, 2002; Clark, 1997). We can gain a better understanding of how people interact with their surroundings and the meanings they assign to items by considering concepts as linked to actions. The integration of sensory, motor, and affective processes challenges the idea of "direct perception" as a simple or

unmediated experience, highlighting the complexity of how we make sense of the world around us.

Challenge for the Zen Perception

Thus, perception is essentially a mediated process that is influenced by action, sensory input, and contextual circumstances. This emphasises that perception is a dynamic interaction with the environment rather than just a single cognitive function. This insight contradicts conventional wisdom, such as that found in Zen Buddhism, which promotes seeing the world as it is, devoid of conceptual illusions. This poses an intriguing query: Is it possible for a Zen master to truly see without the need of intermediaries? If a Zen master could perceive reality as it is, would they not notice the Müller-Lyer illusion?

Zen Perception

Mindfulness, which has its roots in the idea of "bare attention," is emphasised in Zen Buddhism. An unmediated perception of reality, which enables people to witness their experiences free from commentary, analysis, judgment, or conclusions, is what distinguishes this type of consciousness. The core of Zen practice is mindfulness, which is defined as "just seeing" what is present by D.T. Suzuki (2002). By encouraging practitioners to perceive the environment directly, this bare attention approach cultivates a heightened awareness of the present moment free from the distractions of preconceived notions and cognitive processing. This method differs from traditional cognitive theories in the context of perception, which frequently see perception as a process that is fraught with judgments and interpretations.

However, the concept of "bare attention" in Zen Buddhism, which emphasises an undistracted awareness of immediate sensory experiences (Suzuki, 2002), can be likened to the notion of early vision as discussed earlier in this paper in reference to Raftopoulos. "Bare attention" involves observing phenomena without the interference of judgment, interpretation, or cognitive elaboration, allowing one to experience reality in its raw form, and so this may align with Raftopoulos' discussion of early vision, where perception is seen as a direct, immediate response to sensory stimuli before the layers of meaning and interpretation are added. Raftopoulos posits that early vision serves as a foundation for more complex forms of perception, providing a baseline from which meaningful interpretation can arise. However, while this initial sensory awareness is crucial, it may not suffice for meaningful engagement in everyday situations. This type of perception may fall short of generating meaningful interpretations in the context of everyday experiences. Although "bare attention" enables perception free

from cognitive clutter, it does not result in a complete comprehension of objects in a real-world, lived context. On the other hand, meaningful perception is not solely the result of propositional, conscious, or intentional interpretations. It results from a complicated interaction between sensory data and culturally specific behaviours that influence how we engage with the outside environment. The embodied cognition perspective holds that "enactive orientations" toward objects—a dynamic adaptation to these objects in the quest for effective action—are what shape our perceptions (Varela et al. 1991). This adaptive involvement shows that perception is intrinsically connected to our activities and goals in our surroundings and is not merely a passive reception of inputs.

Moreover, this enactive approach underscores that perception is deeply rooted in the culturally shared purposes and goals that inform our interactions. As Wittgenstein (1953) suggests, our understanding of objects is inherently tied to our "forms of life"—the various social practices and contexts in which we are embedded. Therefore, meaningful perception is characterised by a thorough-going embodied-enactive-embedded process, whereby our experiences of objects are shaped not just by sensory input but also by our history, culture, and the actions we take in response to those inputs.

While aiming for a state of non-attachment, even a Zen master maintains a basic engagement with the environment, where meanings are developed via interaction and adaptation. This dynamic link is consistent with Heidegger's idea of Dasein, or being-in-the-world, which emphasises that perception is an active engagement with a contextually rich environment rather than a passive reception of sensory input (Heidegger, 1962). Merleau-Ponty (1962) goes on to say that our embodied interactions are the source of our knowledge, implying that lived experiences rather than static object meanings are what give them their meanings. Thus, even for a Zen master practising mindfulness and aiming for "bare attention", they will still interpret the world through the lens of their past experiences and cultural contexts, engaging with objects in ways that reflect their understanding of the interplay between perception and action.

The Zen practice of non-attachment does not imply a disengagement from the world but rather a deeper awareness of the moment, free from preconceived judgments or desires (Suzuki, 2002). This "bare attention" allows for a clearer perception of reality, yet it is crucial to recognise that such awareness occurs within a framework of prior interactions and meanings shaped by culture and context. Therefore, even as a Zen master seeks to see things as they are, they remain a Dasein whose perception is inherently situated, influenced by their experiences and relationships. This perspective acknowledges that meanings arise not solely from internal cognition or

external rules but from the rich, dynamic interplay between the perceiver and the perceived (Ingold, 2000; Varela et al., 1991).

What is Zen? In essence, it is to attain the stillness of mind under a kind of enlightenment called "Satori", which involves getting free of encumbrances. The Zen state of mind is expansive. It goes beyond the conventional mental constructions and makes one see the things 'as they are' by direct seeing in a childlike innocence, and by going beyond the linguistic, conceptual verbalisations (Fromm, Suzuki, Martino, 1960, pp. 95, 113, 121). Its methods include meditation (achieving a still mind beyond all conceptualisations, cultural expectations and conditionings, etc. It also prescribes the usage of *Koans* to make one go beyond the contradictory concepts in a single thought about something and understand the reality in an unfettered manner, directly without being troubled by the apparent/real contradictions in linguistic constructions. In Zen mind one doesn't feel hatred or cravings because ultimately what we hate or crave is just a mental construction.

Richard Davidson and his associates have done pioneering research on the cognitive and emotional effects of meditation, where they have focused extensively on Buddhist monks trained in meditation practices, particularly those from the Dalai Lama's monastery. These monks, having accumulated over 10,000 hours of meditation training, have shown remarkable enhancements in cognitive capacities that extend beyond what is typically observed in non-meditators. For instance, Davidson and his colleagues found that these monks possess superior perceptual abilities, enabling them to identify hidden images in rapidly changing visual streams—tasks that other subjects, lacking such extensive meditation experience, often find challenging (Davidson, 2003). This implies that regular meditation practice may improve perceptual awareness and attentional control.

Research shows that mindfulness meditation may improve attentional control, emotional regulation, and perceptual clarity (Lutz et al. 2007). By focusing on the present moment with no bias, practitioners of mindfulness are able to observe things objectively and without the addition of personal prejudices. According to studies, practising mindfulness on a regular basis can reduce habitual thought patterns. The practice of Mindfulness involves keeping a non-judgmental awareness of the present moment. This helps the practitioners see things as they are without involvement of mental constructs and subjective biases. It has also been found in various research that regular mindfulness practice can lead to a reduction in habitual thought patterns, which is akin to the Zen perception that prescribes seeing things as they are without distortion by the encumbrances of hatred, greed, competition, anxiety, etc.

Studies have also explored how meditation can enhance sensory perception. Empirical studies conducted by Zhong et al. (2024) reported that participants who engaged in focused attention meditation proved to be more sensitive to the chosen auditory or visual stimuli compared to those who did not meditate. The study says, "*... meditation can positively impact the attentional assignment of a single sensory path is supported by these findings*." Similarly, practices like Samatha and Vipassana have been shown to improve visual and tactile acuity, suggesting that meditation can refine the perceptual process by enhancing focus and reducing noise (Kerr et al., 2011).

In this vein, Buddhism, not believing in anything permanent and seeing things existing as a result of contingent mental constructions, believes merely in the Nirvikalpa ("indeterminate") perception. Certain empirical studies on Buddhist monks, who are also experienced meditators for decades, demonstrate remarkable perceptual abilities in consonance with such metaphysical belief, e.g. Slagter et al. (2007) studies show that they could see certain visual items (alphabetical letters or numbers) embedded inside a very narrow temporal window, in the scale of milliseconds (ms), under what is called an Attentional-Blink Task. On every trial, between 15 and 19 items were presented at the centre of the screen, preceded by a 1780-ms fixation cross. Most of the items were letters, presented for 50 ms each and followed by a 34-ms blank. On T2-present trials, there were two target numbers (T1 and T2) among the items, which participants had to detect and report at the end of the trial. The temporal distance between T1 and T2 could be short (336 ms) or long (672 ms). The following diagram in this connection is helpful to understand this phenomenon.

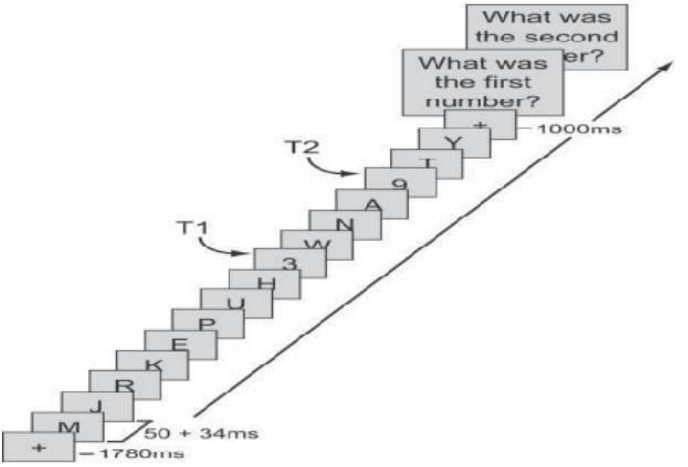

Figure 4: Attentional-Blink Task (Slagter et al., 2007)

These findings seem to validate the claims of Zen that emphasise the cultivation of attentive, non-dual, and non-judgmental awareness as a pathway to authentic or 'true' perception. However, it can be argued that when a person looks at the world for certain objects, he/she already have a goal or purpose in mind and are equipped with certain conceptual and background knowledge about the world, and are aware of what kind of use those objects can be put to, or what impact or influence they can have on him/her. Thus, what one sees in the world is not Nirvikalpa or indeterminate perception; it is always some meaningful perception of the object. It is *enactive*, i.e., oriented towards some action and is mediated with the action-oriented purposefulness. A table and a door, e.g., both made up of the same material, i.e. wood, yet have different meanings for humans for their different uses or actions they can be put into. Similarly, the Muller-Lyer illusion, as we have seen, doesn't exist without a reason; it is a mark of acclimatisation of the person in his/her physical environment, and shows that one is interpreting the world in the right manner. Thus, conceptually, a Zen master, though free from certain affective encumbrances, still has to read the physical world in the right manner, and has to see the Müller-Lyer illusion as is seen by other ordinary persons. Seeing this illusion is not a symbol of 'error', but shows his/her ability to understand the physical world well for basic survival.

Conclusion

This paper proposes that a Zen master may not perceive the "Tent as a Bear" phenomenon because they are less anxious or do not have worry at all. Therefore, they must modulate the repetitive processes that create the final perceptual object. The "Tent as Bear" phenomenon is an excellent example of how contextual elements like anxiety and cognitive framing may affect perception. Because of the brain's top-down processing methods and increased anxiety, regular people might mistake a tent for a bear in this situation. According to Predictive Processing Theory, this occurrence is consistent with the "error-minimising hypothesis," which holds that people interpret ambiguous information by using contextual clues and past knowledge. Thus, when faced with potential threats, ordinary perceivers may be more likely to misinterpret innocuous objects, like a tent, as dangerous, leading to heightened anxiety and misperception. However, when considering the perspective of a Zen master or practitioner, it is essential to recognise their unique cognitive and perceptual processes. Although they may experience reduced anxiety, this does not imply immunity to perceptual illusions such as the Müller-Lyer illusion. Zen practitioners cultivate a state of awareness that emphasises non-attachment and direct experience; yet, they remain enactive beings who interpret the world based on lived experiences and contextual information.

How, then, does this apply to the Müller-Lyer illusion? Does a Zen master or a Zen practitioner see the Müller-Lyer illusion? I predict – yes. There can be no immunity even to the Zen Mind's perception since even a Zen master is an enactive being, for even he/she needs to read the world in a well-adjusted or acclimatised manner. This perspective aligns with the principles of enactive cognition, which posits that perception is inherently tied to the agent's interactions with their environment. Zen practitioners, despite their advanced meditation practices and efforts to cultivate mindfulness and non-attachment, must still navigate a world filled with sensory

information that requires interpretation. As they perceive and interact with their surroundings, they draw upon contextual knowledge and experience to make sense of ambiguous stimuli. Thus, even a Zen master perceives the world through a lens shaped by prior experiences, cultural understanding, and situational context.

Endnotes

[1] They claim (P. 236): "…we offer the view that person schemata are essentially based on direct perception while person images are essentially relying on the interpretation of situations involving background knowledge and the construction of narratives. As we have emphasised in our criticism of Goldman's Simulation-Theory, we have to acknowledge various quite different strategies of understanding others. We have distinguished, broadly, non-conceptual from conceptual understanding. Which of these strategies is (or needs to be) employed in order to understand another person depends crucially (a) on the person in question and our prior relation to and familiarity with that person (that is, basically, on the richness of our person-model regarding that person), (b) on the situation and context and, finally, (c) on the type and complexity of the mental state(s) in question. All these three dimensions have to be taken into account in developing a persuasive theory of understanding other minds."

References

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature, 221*(5184), 963–964. https://doi.org/10.1038/221963a0

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. MIT Press.

Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Santorelli, S. F., Sheridan, J. F., & Urbanowski, F. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic Medicine, 65*(4), 564–570. https://doi.org/10.1097/01.PSY.0000077505.67574.E3

Davidson, R. J., & Lutz, A. (2008). Buddha's brain: Neuroplasticity and meditation. *IEEE Signal Processing Magazine, 25*(1), 174–176.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.

Fodor, J. A. (1985). Précis of *The modularity of mind*. *Behavioral and Brain Sciences, 8*(1), 1–5. https://doi.org/10.1017/S0140525X0001921X

Fromm, E., Suzuki, D. T., & De Martino, R. (1960). *Zen Buddhism and psychoanalysis*. Harper & Row.

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

Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology, 56*(5), 872–876. https://doi.org/10.1037/h0040 546

Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Harper & Row. (Original work published 1927)

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hutto, D. D. (2005). Knowing what? Radical versus conservative enactivism. *Phenomenology and the Cognitive Sciences, 4*(4), 389–405. https://doi.org/10.1007/s11097-005-9001-z

Ingold, T. (2021). *The perception of the environment: Essays on livelihood, dwelling and skill*. Routledge.

Ingold, T., & Kurttila, T. (2000). Perceiving the environment in Finnish Lapland. *Body & Society, 6*(3–4), 183–196. https://doi.org/10.1177/1357034X00006003010

Juckel, G., Heinisch, C., Welpinghus, A., & Brüne, M. (2018). Understanding another person's emotions—An interdisciplinary research approach. *Frontiers in Psychiatry, 9*, Article 414. https://doi.org/10.3389/fpsyt.2018.00414

Kerr, C. E., Jones, S. R., Wan, Q., Pritchett, D. L., Wasserman, R. H., Wexler, A., & Moore, C. I. (2011). Effects of mindfulness meditation training on anticipatory alpha modulation in primary somatosensory cortex. *Brain Research Bulletin, 85*(3–4), 96–103. https://doi.org/10.1016 j.brainresbull.2011.03.026

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*(3), 389–412. https://doi.org/10.1016/0042-6989(94)00176-M

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology, 6*(3), 547–569 https://doi.org/10.1007/s13164-015-0253-4

Lutz, A., Dunne, J. D., & Davidson, R. J. (2007). Meditation and the neuroscience of consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 499–555). Cambridge University Press.

Merleau-Ponty, M. (1962). *Phenomenology of perception* (C. Smith, Trans.). Routledge & Kegan Paul. (Original work published 1945)

Newen, A., & Schlicht, T. (2009). Understanding other minds: A criticism of Goldman's simulation theory and an outline of the person model theory. *Grazer Philosophische Studien, 79*(1), 207–242.

Newen, A., & Vetter, P. (2017). Why cognitive penetration of our perceptual experience is still the most plausible account. *Consciousness and Cognition, 47*, 26–37. https://doi.org/10.1016/j.concog.2016.09.005

Noë, A. (2004). *Action in perception*. MIT Press.

Nolen-Hoeksema, S., Fredrickson, B. L., Loftus, G. R., & Wagenaar, W. A. (2009). *Atkinson & Hilgard's introduction to psychology* (15th ed.). Wadsw orth Cengage Learning.

Pedersen, D. M., & Wheeler, J. (1983). The Müller-Lyer illusion among Navajos. *The Journal of Social Psychology, 121*(1), 3–6. https://doi.org/10.1080/00224545.1983.9924459

Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). Routledge & Kegan Paul.

Prinz, J. J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. MIT Press.

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. MIT Press.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(3), 341–365. https://doi.org/10.1017/S0140525X99002022

Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. MIT Press.

Raftopoulos, A. (2001). Is perception informationally encapsulated? The issue of the theory-ladenness of perception. *Cognitive Science, 25*(3), 423–451. https://doi.org/10.1207/s15516709cog2503_4

Raftopoulos, A. (2009). *Cognition and perception: How do psychology and neuroscience inform philosophy?* MIT Press.

Raftopoulos, A. (2014). The cognitive impenetrability of early vision is a necessary and sufficient condition for purely nonconceptual content. *Philosophical Psychology, 27*(5), 601–620. https://doi.org/10.1080/09515089.2012.729486

Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1966). *The influence of culture on visual perception*. Bobbs-Merrill.

Slagter, H. A., Lutz, A., Greischar, L. L., Nieuwenhuis, S., Davidson, R. J., & Kastner, S. (2007). Mental training affects distribution of limited brain resources. *PLoS Biology, 5*(6), Article e138. https://doi.org/10.1371/journal.pbio.0050138

Suzuki, D. T. (2002). *Buddha of infinite light: The teachings of Shin Buddhism, the Japanese way of wisdom and compassion*. Shambhala Publications.

Suzuki, S. (2020). *Zen mind, beginner's mind*. Shambhala Publications.

Taylor, C. (2007). Engaged agency and background in Heidegger. In C. B. Guignon (Ed.), *The Cambridge companion to Heidegger* (2nd ed., pp. 202–221). Cambridge University Press.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Blackwell.

Zhong, S.-Y., Guo, J.-H., Zhou, X.-N., Liu, J.-L., & Jiang, C.-L. (2024). Effects of brief mindfulness meditation training on attention and dispositional mindfulness in young adult males. Acta Psychologica, 246, 104277. https://doi.org/10.1016/j.actpsy.2024.104277