An integrative model of visual control of action

Um modelo integrativo do controle visual da ação

Un modelo integrador de lo control visual de la acción

Elton Hiroshi Matsushima[1]

[1] Universidade Federal Fluminense – UFF. *E-mail*: eh.matsushima@gmail.com *ORCID*: https://orcid.org/0000-0003-3392-4632

Informações do artigo:

Elton Hiroshi Matsushima
eh.matsushima@gmail.com

Recebido em: 29/04/2020
Aceito em: 15/07/2020

ABSTRACT

The present study offers an integrative proposal of a model of visual control of action, specifically in relation to the visually directed tasks. Within these models, the calibration between visual signals and vestibulo-kinesthetic signals is of fundamental importance, especially in the case of visually directed tasks. The Hierarchical Control Model (Marken, 1985), the Functional Organization Model (Rieser et al., 1995), the Time-based Heuristics (Lederman et al., 1987), and the Model of Visual Control of Locomotion (Lee & Lishman, 1977b), are integrated into a single model, which still incorporates recent developments in empirical research. The proposed model provides a theoretical framework to guide the experimental research of the visual control of action, in order to determine the processing steps and paths not yet clarified by the empirical evidence.

KEYWORDS:

Visual Perception; Spatial Perception; Perceptual-Motor Processes; Motor Control.

RESUMO

O presente estudo apresenta uma proposta integrativa dos modelos de controle visual da ação, principalmente no que diz respeito às tarefas visualmente dirigidas. Dentro destes modelos, a calibração entre os sinais visuais e os sinais vestibulo-cinestésicos é de fundamental importância, especialmente no caso das tarefas visualmente dirigidas. O Modelo de Controle Hierarquizado (Marken, 1985), o Modelo de Organização Funcional (Rieser et al., 1995), a Heurística Temporal (Lederman et al., 1987) e o Modelo de Controle Visual da Locomoção (Lee & Lishman, 1977b), são integrados dentro de um único modelo, que ainda incorpora desenvolvimentos recentes da pesquisa empírica. O modelo proposto fornece um arcabouço teórico para orientar a pesquisa experimental do controle visual da ação, de forma a determinar as etapas e os fluxos processuais ainda não esclarecidos pelas evidências empíricas.

PALAVRAS-CHAVE:

Percepção Visual; Percepção Espacial; Processos Perceptomotores; Controle Motor.

RESUMEN

El presente estudio presenta una propuesta integradora de los modelos de control visual de la acción, principalmente con respecto a las tareas dirigidas visualmente. Dentro de estos modelos, la calibración entre señales visuales y señales vestibulo-kinestésicas es de fundamental importancia, especialmente en el caso de tareas dirigidas visualmente. El Modelo de Control Jerárquico (Marken, 1985), el Modelo de Organización Funcional (Rieser et al., 1995), la Heurística Temporal (Lederman et al., 1987), y el Modelo de Control de Locomoción Visual (Lee & Lishman, 1977b) se integran en un solo modelo, que todavía incorpora desarrollos recientes en la investigación empírica. El modelo propuesto proporciona un marco teórico para guiar la investigación experimental del control visual de la acción, con el fin de determinar los pasos y flujos de procesamiento aún no aclarados por la evidencia empírica.

PALABRAS CLAVE:

Percepción Visual; Percepción Espacial; Procesos Perceptuales-Motores; Control Motor.

The final part of the last century observed the onset and development of a new concept in the visual space perception assessment, the visually directed tasks, named after the task created by Foley and Held (1972), *visually directed pointing*. The concept was conceived in opposition to the *visually guided tasks*, which were tasks accomplished under continuous visual monitoring of performance. On the contrary, visually directed tasks were accomplished under intermittent visual monitoring of performance. In most of the experimental settings, these tasks were accomplished with vision occluded towards previously seen targets (Loomis, Da Silva, Fujita, & Fukusima, 1992).

Although the study of this new form of visual perception measurement has been going for several years now, findings are not unequivocal. Some studies evidenced accuracy in those visually directed tasks, in locomotion (Bigel & Ellard, 2000; Fukusima, Loomis, & Da Silva, 1997; Loomis et al., 1992; Matsushima, Ribeiro-Filho, Douchkin, & Da Silva, 2002; Muroi & Higuchi, 2017; Philbeck & Loomis, 1997; Rieser, Ashmead, Talor, & Youngquist, 1990; Thomson, 1983), throwing (Thomson, 1983), and pointing (Foley & Held, 1972; Fukusima et al., 1997), while others showed systematic errors in tasks, such as locomotion (Elliott, 1986, 1987; Matsushima, Gomes, Ribeiro-Filho, & Da Silva, 2001; Steenhuis & Goodale, 1988; Sun, Campos, Young, Chan, & Ellard, 2004), negotiating an aperture (Muroi & Higuchi, 2017), and throwing (Eby & Loomis, 1987). One may consider this controversy secondary when considering that the common finding was a more than adequate performance in providing observers with adaptive behaviors to varied environmental conditions.

The central issue of the present study is related to the processing of visual information needed to provide those adequate performances in visually directed tasks, specifically locomotion. Loomis and colleagues (1992) described the constituent subprocesses to accomplish this kind of task: a) visual perception of the spatial location of target and other landmarks; b) updating of present self-position based on idiothetic information (e.g., perceived self-velocity, Glasauer, Amorim, Viaud-Delmon, & Berthoz, 2002); c) imagined updating of the target spatial location, based on updated self-position; and d) execution of appropriate response towards the target's updated spatial location. These subprocesses are essential to any visually directed task and the errors and variabilities associated in those tasks are due to errors and variabilities accumulated across these different subprocesses.

Loomis and Philbeck (2008) proposed a theoretical model to describe the relation between sources of information and visually directed tasks (Figure 1). In this model, a single internal variable, the perceived location, was the input to any task related to visual distance perception, both perceptual (e.g., verbal reports) and motoric (e.g., visually directed tasks). Results of several experimental investigations support this model with strict covariation of performances under those response modalities as a function of availability of the

visual cues (Philbeck & Loomis, 1997; Philbeck, Loomis, & Beall, 1997; Matsushima et al., 2002; Matsushima, Chiaretti, Kreling, Lima, Da Silva, & Ribeiro-Filho, 2004; Sun et al., 2004).
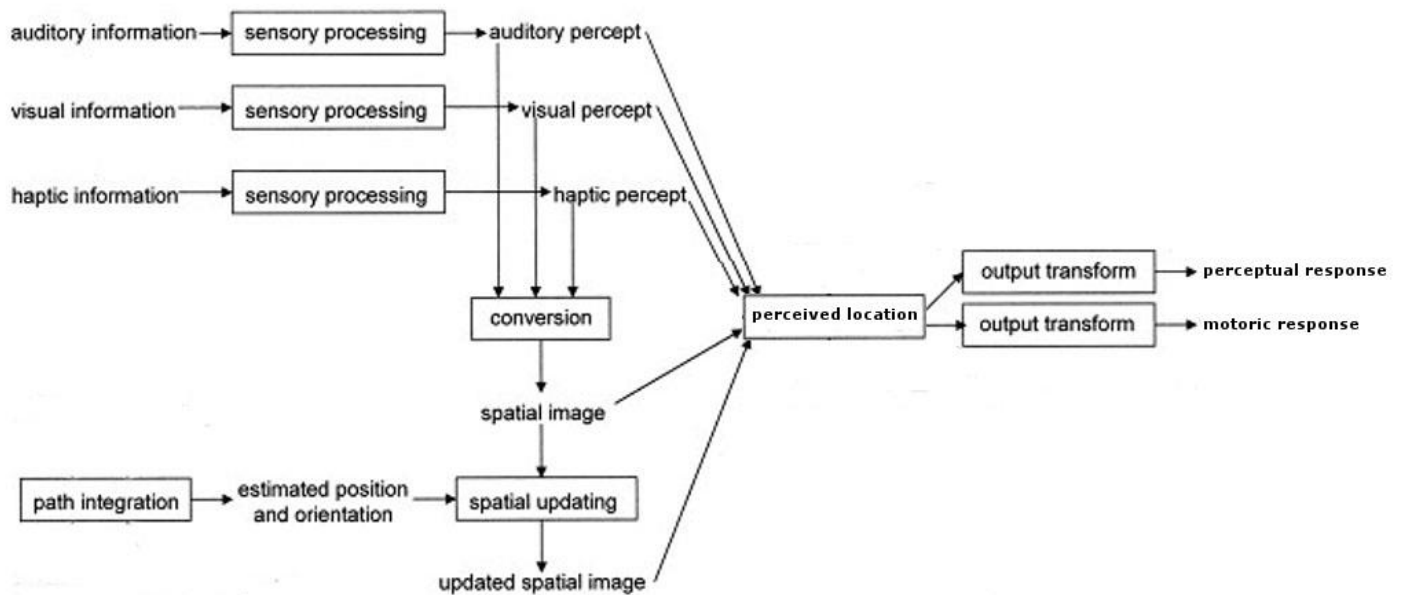
Source: Own elaboration.

Figure 1.

_____

Loomis and Philbeck (2008) model of perceptually directed tasks.

     A consequence of the above-cited model is that performances in both behavioral responses, verbal reports and motoric responses, would be intercorrelated. The single internal variable, perceived spatial location, produced by the combination of sources of information, underwent simple transform processes to feed into action and perception systems. These specific transforms may be related to neurophysiological findings supporting the dissociation between neural systems of visual processing.

     These neurophysiological evidences suggest that visual information undergoes separate processing in two neural pathways, the ventral pathway (VP) and dorsal pathway (DP), which connect occipital lobes to two different other cortical areas, inferotemporal cortex (IT) and posterior parietal cortex (PP), respectively (Goodale & Milner, 1992; Ungerleider & Mishkin, 1982). The functional differences associated to those visual processing pathways were determined in a series of studies run by Goodale and Milner (1992), whose

results showed that the system composed by DP and PP processes action-specific information, such as structural features and object orientation, while the VP-IT system would be responsible for object recognition, processing detailed and distinctive object features, such as shape, color and surface.

Despite the unequivocal evidence for two parallel visual pathways, one must not confound parallel processing with independency between systems (Matsushima, 2004; Similä & McIntosh, 2015). Other research teams (Bruno, 2001; Feldman, 2009; Vishton, Rea, Cutting, & Nuñez, 1999) suggested that instead of functional separation, different frames of reference would have produced the differences found between visual pathways. A frame of reference can be defined as a locus or a set of loci to which spatial locations were determined. They may be either egocentric or allocentric. In an egocentric frame of reference, these loci defining spatial locations were all located in observer's body, i.e., an observer-centered frame of reference, based on polar coordinates whose origin is the observer location. In an allocentric frame of reference, spatial locations were defined by external loci, thus a coordinate system whose origin may be in objects or external landmarks (Matsushima & Ribeiro-Filho, 2003). The DP would be responsible to visual processing in an egocentric frame of reference, a perceptual coding appropriate to action-oriented visual processing (for neurophysiological evidence, see Andersen, Snyder, Bradley, & Xing, 1997; Andersen & Buneo, 2002). The VP would be responsible for visual processing in an allocentric frame of reference, more suitable to coding object features and its spatial relations. There are still others that suggest that neural segregation would represent a specialization in proximal (at arm reach) and distal spaces, instead of action and object recognition specializations. In this view, DP would be dedicated to proximal space and VP to distal space (Weiss, Marshall, Zilles, & Fink, 2003).

Another alternative view would be that entailed by the Visual Attention Model (VAM, Schneider, 1995). It proposes that vision-for-perception and vision-for-action share an early selection mechanism, before the splitting in ventral and dorsal pathways. This selection defines an object token corresponding to the target, in which perceptual features are transmitted to VP, whilst information about its location and spatial features is forwarded to DP (Similä & McIntosh, 2015).

Considering the experimental findings, one must conceive the visual system as an interactive set of processing modules, composed by two parallel neural pathways, richly interconnected. DP would be a pathway processing proximal space (at arm's reach) by an implicit manipulation of visual information in an egocentric frame of reference, richly connected to motor cortices. VP would be related to detailed analysis of visual information, richly connected to high-order cortices and mesencephalic structures (dedicated to high order cognitive skills, such as memory and thinking). Quotidien tasks usually recruit both pathways working in synergistic ways, while independent functioning would only happen under extreme conditions, such as cerebral lesions or information-deprived conditions (Matsushima, 2004).

These two pathways would receive the single internal variable, perceived location, and proceed to transform it to their specific goals resulting in the different behaviors respectively associated with these pathways. Considering the sequence of subprocesses of a visually directed task, the perceived location matches the first processing step, visual perception of the target. This subprocess would be updated by integration to the updated self-position (on its turn, updated by perceived self-velocity). This updating process would be accomplished by the global exproprioceptive system (Lee & Lishman, 1977a), composed by different perceptual systems, which include visual, vestibular and kinesthetic systems. Environments thrive in information available for spatial updating, thus observers should easily update their position and orientation while moving in space, even without vision (Avraamides, Klatzky, Loomis, & Golledge, 2004). Evidence also suggests that vestibular and kinesthetic information (the latter from muscle spindles) can support accurate updating of spatial orientation, since accurate performance has been demonstrated in path integration during passive transport, based on sustained spatial tuning activity of hippocampal cells (Lackner & DiZio, 2005).

Exproprioception system thus integrates the self-direction and self-velocity (proprioception) with self-location (exteroception) (Lee & Lishman, 1977a). The processes accomplished by exproprioceptive system can be conceived as the synergistic calibration existing between perceptual and action systems (Rieser, Pick, Jr, Ashmead, & Garing, 1995). Living organisms execute actions towards specific adaptive goals, or target-directed actions, consisting of sets of perception-action couplings, i.e., relations between certain egocentric locations and

actions directed towards those locations. In other words, to each perceived egocentric location, there would be a definite set of properties of possible actions. However, the correspondence between perceived egocentric location and certain action properties would not be a fixed set, considering that the magnitude and direction of actions would depend on conditions subject to time and environmental variables. Therefore, the calibration is a flexible process, task-dependent, environment-dependent (Redding & Wallace, 2001; Rieser et al., 1995), and also dependent on body dynamics changes (Leclere, Sarlegna, Coello, & Bourdin, 2019).

Several calibration models were proposed in the scientific literature. A relevant one was Time-based Heuristics (Lederman, Klatzky, Collins, & Wardell, 1987), which states that, when step-counting was suppressed, observers combine temporal units to perceived self-velocity to compute the walked distance. This model explains the updating of self-position without continuous visual monitoring, accomplished by integration of perceived self-velocity and temporal units elapsed into an updated self-position.

In the Hierarchical Control Model (Marken, 1986), coordinated actions, those that require several responses to achieve a planned goal, are accomplished by a production system. In a basic production system, the sensor transforms the environment stimulus into a perceptual signal which feeds up the high-order level and is compared in a comparison unit with the reference copy (goal). In its turn, comparison unit transforms the difference between perceptual signal and reference copy in an error signal, which feeds into the effector to transform it in corrective response. Analyzing the system overall, it is not dedicated to produce outputs (responses), but to produce sensorial inputs which would be equal to reference copy, i.e., to alter the environment or organism status until the new perceptual layout matches the reference copy of goals. The calibration of this system would be online, i.e., would need continuous monitoring of both environment and body status. To a visually directed task, the system would work without one of the sensorial subsystems, visual system, while using the remaining subsystems, auditory, vestibular, kinesthetic and motoric ones.

According to experimental evidence, humans can apply time-based heuristics to visually guided tasks, such as locomotor pointing (de Rugy, Montagne, Buekers, & Laurent, 2002), since the useful information for this type of task is *tau,* or *time-to-contact* (Lee, 1976). On the other hand, for visually directed tasks, as the

updating of self-position in walking tasks with vision occluded, the evidence was not unequivocal (Lederman et al., 1987; Schwartz, 1999) and suggests the existence of a body-dimension factor, which could be step length, as the fundamental element for the computing of walked distance (Schwartz, 1999).

In the Functional Organization Model (Rieser et al., 1995), the fine-tuning of a behavior by the effect of practice would generalize to any other behaviors towards the same targets or goals. In this model, the calibration would fine-tune at once all methods of achieving a certain spatial goal to accommodate the changes of environmental and body conditions. For human locomotion, it has been demonstrated that the calibration system is responsive to changes in the covariation of biomechanical activity and optical flows.

In the Model of Visual Control of Locomotion (Lee & Lishman, 1977b), there are three levels of control: 1) a superior level, highly cognitive, identified with the formulation of a general plan for accomplishing tasks; 2) an intermediate level, which involves the perception of the spatial layout of observer's surroundings to determine the path to be followed and the execution of this more detailed plan; and 3) a lower level, which operates the regulation of speed, direction and biomechanical aspects to maintain the selected path. The entire calibration process here would take place at the two lowest levels, adapting the path and the biomechanical aspects to possible obstacles and small environmental variations, respectively.

Comparing the other possible models of visuomotor calibration, the Visual Control of Locomotion Model (Lee & Lishman, 1977b) was very general and can be adapted to the more detailed Hierarchical Control Model (Marken, 1986), as well as the indications of Functional Organization Model (Rieser et al., 1995), and the suggestions of two other cognitive representations that should be considered by any visuomotor calibration system: the spatial envelope and the plant dynamics (Loomis & Beall, 1998). The spatial envelope refers to the dimensions and equipment associated with the body, and is also subject to recalibration, due to the bodily changes that can occur, both due to body development and ephemeral restrictions. The plant dynamics refers to restrictions inherent to the constituent structure of observer's body, which constrain the possible transformations of the relative position of limbs and the kinematics caused by the interactions between muscular and gravitational forces.

The Hierarchical Visuomotor Control Model

The present study proposes an integrative model, converging the previously cited models and experimental evidences, the Hierarchical Visuomotor Control Model (Figure 2). The High-Order Control Level determines a general motor plan and can be conceived as high-order cognitive processes, such as planning, reasoning and decision-making systems. This general motor plan will be integrated with information from the Spatial Envelope, Plant Dynamics, Spatial Layout and Self-Position in the Intermediate Control Level to select a more suited path to the current body and environmental restrictions. One may conceive this Intermediate Control Level as the Working Memory, since it integrates long-term memory information (Spatial Envelope and Plant Dynamics) and present information (Spatial Layout and Self-Position). Long-term memory information is integrated by the Episodic Buffer (Baddeley, 2012) along with present spatial location to feed into Central Executive which determines the selected path based on long-term body information and present self-position. Episodic Buffer also works as temporary storage (Baddeley, 2012) for target location (reference copy or general motor plan). Central Executive perform cognitive manipulation of information from slave subsystems, Episodic Buffer and Visuospatial Sketchpad (Baddeley, 2012), to compare updated self-position with the remembered target position.
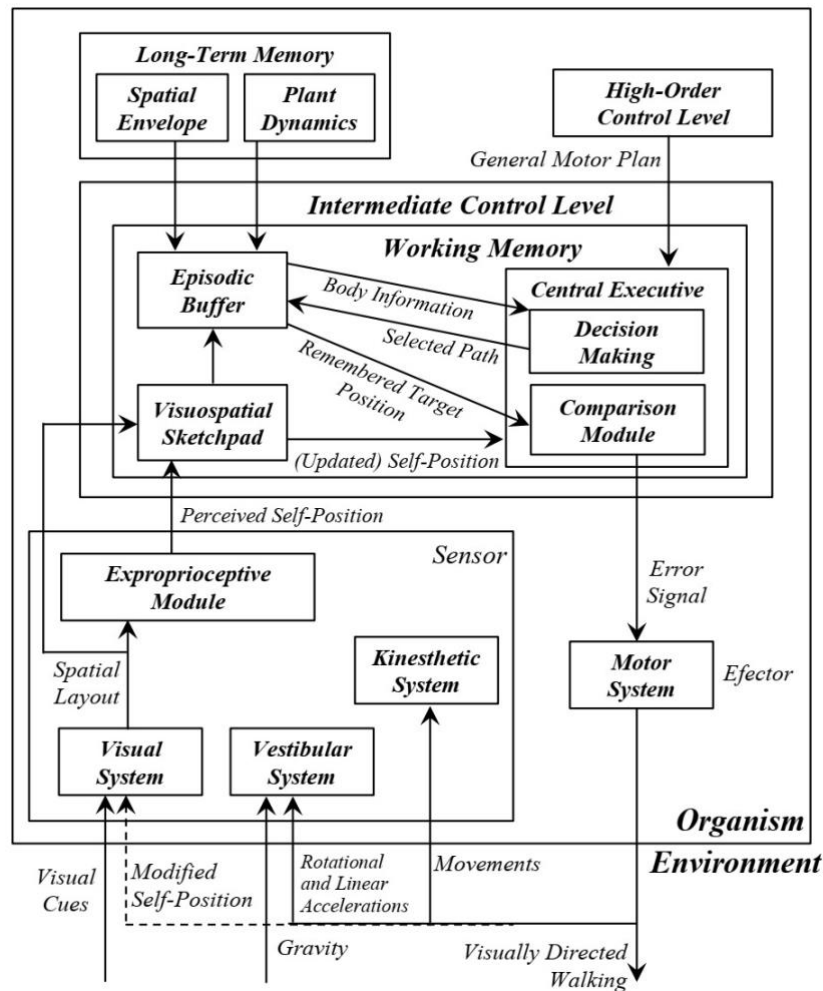
Source: Own elaboration.

Figure 2.

Model of Hierarchical Visuomotor Control (see text for details and descriptions).

The selected path is the reference signal that the central executive will compare with the perceptual signal produced by the perceptual systems (visual, kinesthetic and vestibular systems), and the difference found will be transformed into an error signal for the motor system to generate a corrective action. This action will generate movements and accelerations in the body that will feed-back into kinesthetic and vestibular systems, respectively (Lackner & DiZio, 2005), and will also generate changes in self-position, which will feed-back into visual system. These perceptual systems will have their signals integrated, through a hypothetical integrating module, Exproprioceptive Module, generating a new perceptual signal (updated

spatial location in a visually directed task), which will be compared to the reference signal (remembered spatial). Modified self-position, a consequence of the visually directed walking, indirectly feeding into the visual system, since vision is occluded in those tasks, was represented by a dashed arrow. In a visually guided task, the arrow should be continuous, representing a direct access.

This model integrates experimental findings with some hypotheses suggested by several researchers. The signals resulting from an observer's displacement over the visual system are intrinsically linked to the signals resulting from this same displacement over the vestibular and kinesthetic systems. Previous research has shown that this connection can be disturbed by the imposition of a walking speed different from that of optical flows (Rieser et al., 1995) or by the introduction of conflicts between visual and vestibulo-kinesthetic information (Ellard & Shaughnessy, 2003). If the relationship between visual signal and vestibulo-kinesthetic signals, learned throughout the observer's development, is broken by the introduction of a "correction" factor, the responses will tend to converge to a midpoint between the response based on the long learning of development and the "corrected" response in a recalibration (Ellard & Shaughnessy, 2003).

Neurophysiological evidence suggests a separation between motor control and spatial updating processes (Pierce, Saj, & Vuilleumier, 2019). Posterior parietal areas might hold a sensory representation of visual space primarily based on bottom-up inputs (visual – Goodale & Milner, 1992; Ungerleider & Mishkin, 1982; vestibular and kinesthetic – Lackner & DiZio, 2005), thus related to spatial updating processes, while more anterior parietal areas might integrate these representations with top-down factors related to attention and motor demands of the task, which would be more influenced by high-order cognitive processes. This evidence was instantiated in separate modules in Central Executive performing a high-order decision-making and a simpler comparison module of perceptual information and stored spatial representation.

The limitation of this model would be that it may not be generalized to online visual control of the action. In online visual control, evidence suggests that action is guided by visual information available during movements. Under some conditions, e.g. reduced-cue conditions or repeated object motion, action may be controlled by heuristics, mappings, or spatial memory, in terms of an allocentric stored representation

(Manzone & Heath, 2018). For instance, in visually directed tasks, a remembered spatial location of the target appears to guide blind walking after vision is removed (Zhao & Warren, 2015). Although, yet to be thoroughly investigated, this assertion must also deal with the fact that it is unlikely that every human behavior must be online controlled, considering our attentional limits. Several evidences showed that human locomotion performance remains unchanged under intermittent visual monitoring (Bigel & Ellard, 2000; Fukusima et al., 1997; Loomis et al., 1992; Matsushima et al., 2002; Muroi & Higuchi, 2017; Philbeck & Loomis, 1997; Rieser et al., 1990; Thomson, 1983). These evidences advocate for an internal perceptual variable underlying action performance, favoring a model-based visual control of action, such as the one proposed here.

Another limitation would be a simplified specification of how visual information would entail motor control. Research on motor control provided detailed descriptions, including neurophysiological pathways, as instantiated in the Threshold Position Control (Feldman, 2009). Our nervous system was able to set the threshold position of any joint, referent limb configuration (muscles of the whole limb) or even referent body configuration (muscles in the whole body), i.e., the position at which muscles are ready to be activated in response to displacements. After the end of a movement, the near-threshold state was restored in the new position, so the nervous system shifts the threshold position when another intentional action is produced. Therefore, we can set a referent position of the whole body in the environment and elicit any goal-directed action by shifting this referent position (for a complete description, see Feldman, 2009).

Threshold position control is related to the notion of frames of reference. The threshold position of referent body configuration can be considered as the origin of a frame of reference comprised of all-possible body configurations represented in this frame of reference (Paillard, 1991). Another frame of reference in this theory represents all-possible positions of the whole body relative to the environment. When we specify a threshold position of the body in the environment, it sets all the distances and directions of the other body postures from the baseline threshold. The contraction-relaxation muscular pattern generates activity based on this difference between its current state and the threshold state in the origin (Feldman, 2009).

Future research

Visually directed tasks would be techniques for accessing visual perception of space more precisely than verbal responses, such as verbal reports and magnitude estimation (Loomis & Beall, 2008). This statement is based on the impenetrability of visually directed actions to cognitive biases, such as the observer's perceptual tendencies (eg, equidistance tendency and specific distance tendency, Gogel, 1965, 1974; Gogel & Da Silva, 1987; Gogel & Tietz, 1973).

The judgments of perceived distance are subject to cognitive influences that can deviate them, thus making it impossible to compare the physical magnitude with the perceived magnitude. This bias was overcome by the head movement technique, outlined by Gogel (1977). He claims that he managed to produce measures of perceived distance free of cognitive influences, which led him to his Bifactorial Theory of Perception of Size and Distance (Gogel & Da Silva, 1987). It establishes two processes occurring in parallel that converge in the judgment of distance and in size, a perceptual and a cognitive process. As the quality and quantity of visual distance information sources are restricted, the more cognitive processes dominate the final convergent result.

These calibration processes raise questions about the validity of visually directed actions as measures of visual perception. Specifically, the calibration would explain the linearity found in some studies between walked distance and physical distance, contrary to systematically compressed errors in perceived distance judgments (Gilinsky, 1951; Künnapas, 1968). Calibration would be the specific factor of response modality, visually directed task, that compensates for this systematically compressed perceptual input converging in the linearity with visually directed actions. Therefore, exploring the limits of calibration in visually directed tasks would lead to a greater understanding of this asymmetry in errors in walked distances and verbally reported perceived distances.

One important issue would be investigating whether the recalibration of a relationship between visual signals (previously perceived spatial location) and vestibulo-kinesthetic signals can be transferred (generalized) to different routes. This would verify one of the open questions regarding the processing of

visual information for visually directed tasks: whether the integration of visual and vestibulo-kinesthetic signals occurs at the second-order level, in the working memory, or at the first-order level, exproprioceptive module. The triangulation task would be used to verify whether the recalibration of the visual signal and vestibulo-kinesthetic ratio affects the perceived distance or the perceived location. If the effect is on the perceived location, walking will be affected in direct walking tasks, as well as in triangulation tasks. If, on the other hand, the effect is only on direct walks, it will only be on the perceived distance.

An issue arising from the model is the working memory role in visually directed tasks. Since working memory processing capacity is limited (Baddeley, 2012) and it has an active role in the model for iterating motor behaviors until achieving the goals, one may consider the overload of working memory would disrupt performance and produce larger errors and variability. This issue arose as well in another study (Ellard & Shaughnessy, 2003). They suggested that target location based on vestibulo-kinesthetic information may decay more rapidly than the one based on visual information, and that visual target location actually interferes with vestibulo-kinesthetic target location. Thus, during a visually directed task, when the visual information is presented after locomotor information, it masks the previous one. This interpretation would imply that working memory, specifically the visuospatial sketchpad would have modality-specific priorities in its processing, probably to avoid overload or competition for its limited processing/storage capacity.

Another research question would be investigating the top-down influences on visually directed tasks. These studies would address questions about the interaction between high-order levels and perceptual levels on visually directed tasks, whether cognitive processes could directly and effectively manipulate spatial representations. In a series of previous studies, decreases in accuracy were found when the cognitive load of visually directed walking tasks was increased. This was accomplished by comparing the performances in a task of blind walking towards imagined avoidance points with the standard performances in a blind walking task. The increase in constant error in the imagination groups suggests that visually directed tasks could also be contaminated by cognitive influences (Matsushima et al., 2001; Gomes, Oliveira, Matsushima, Santos, Ribeiro-Filho, & Da Silva, 1999).

Another issue is related to the functional equivalence between spatial representations created from vision, auditory location and symbolic means, such as maps and language, although there is a cognitive disadvantage in the formation and retention of spatial representations via language (Avraamides, Loomis, Klatzky, & Golledge, 2004). Other results showed an effect of the semantic content of a concurrent evaluation of stimuli in a visually guided task (Olthuis, Van Der Kamp, & Caljouw, 2017), while still others suggest that walking strategy plays an important role in distance accuracy on visually directed tasks during the disturbance of idiothetic information (Glasauer et al., 2002). Experiments that induce overt cognitive manipulation of spatial representations would test hypotheses on this functional equivalence of symbolic spatial representations. Some evidence against this equivalence comes from signal detection theory, which assumes distinct roles of attention and expectation on decision-making. Attention modulates early sensory processing, whereas expectation influences the selection and execution of motor responses. Results support that expectations primarily influence decisions by modulating post-perceptual stages of information processing (Rungratsameetaweemana, Itthipuripat, Salazar, & Serences, 2018). Overt cognitive manipulation would be similar to an expectation on visually directed tasks and thus affect only the post-perceptual stages.

The present study offers a potentially useful model for explaining present research evidences and for inspiring new research, expanding the understanding of visually directed tasks. The Hierarchical Visuomotor Control Model accommodates several previously proposed models, most of the available experimental evidence, and also integrates findings on cognitive processes, such as working memory model, to create a comprehensive explanation of those behaviors.

References

Andersen, R.A., & Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience, 25*, 189-220. https://doi.org/10.1146/annurev.neuro.25.112701.142922

Andersen, R.A., Snyder, L.H., Bradley, D.C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience, 20*, 303-330. https://doi.org/10.1146/annurev.neuro.20.1.303

Avraamides, M.N., Klatzky, R.L., Loomis, J.M., & Golledge, R.G. (2004). Use of cognitive versus perceptual heading during imagined locomotion depends on the response mode. *Psychological Science, 15*(6), 403-408. https://doi.org/10.1111/j.0956-7976.2004.00692.x

Avraamides, M.N., Loomis, J.M., Klatzky, R.L., & Golledge, R.G. (2004). Functional equivalence of spatial representations derived from vision and language: evidence from allocentric judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(4), 804-814. https://doi.org/10.1037/0278-7393.30.4.804

Baddeley, A.D. (2012). Working memory: theories, models, and controversies. *Annual Review Psychology, 63*, 1-29. https://doi.org/10.1146/annurev-psych-120710-100422

Bigel, M.G., & Ellard, C.G. (2000). The contribution of nonvisual information to simple place navigation and distance estimation: an examination of path integration. *Canadian Journal of Experimental Psychology, 54*(3), 172-184. https://doi.org/10.1037/h0087339

Bogler, C., Bode, S., & Haynes, J.-D. (2011). Decoding successive computational stages of saliency processing. *Current Biology, 21*(19), 1667-1671. https://doi.org/10.1016/j.cub.2011.08.039

Bruno, N. (2001). When does action resist visual illusions? *Trends in Cognitive Sciences, 5*(9), 379-382. https://doi.org/10.1016/s1364-6613(00)01725-3

de Rugy, A., Montagne, G., Buekers, M.J., & Laurent, M. (2002). Temporal information for spatially constrained locomotion. *Experimental Brain Research, 146*, 129-141. https://doi.org/10.1007/s00221-002-1155-0

Eby, D.W., & Loomis, J.M. (1987). A study of visually directed throwing in the presence of multiple distance cues. *Perception & Psychophysics, 41*(4), 308-312. https://doi.org/10.3758/BF03208231

Ellard, C.G., & Shaughnessy, S.C. (2003). A comparison of visual and nonvisual sensory inputs to walked distance in a blind-walking task. *Perception, 32*, 567-578. https://doi.org/10.1068/p5041

Elliott, D. (1986). Continuous visual information may be important after all: A failure to replicate Thomson (1983). *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 388-391. https://doi.org/10.1037/0096-1523.12.3.388

Elliott, D. (1987). The influence of walking speed and prior practice on locomotor distance estimation. *Journal of Motor Behavior, 19*(4), 476-485. https://doi.org/10.1080/00222895.1987.10735425

Feldman, A.G. (2009). New insights into action–perception coupling. *Experimental Brain Research, 194*, 39-58. https://doi.org/10.1007/s00221-008-1667-3

Foley, J.M., & Held, R. (1972). Visually directed pointing as a function of target distance, direction, and available cues. *Perception & Psychophysics, 12*(3), 263-268. https://doi.org/10.3758/BF03207201

Fukusima, S.S., Loomis, J.M., & Da Silva, J.A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 86-100. https://doi.org/10.1037//0096-1523.23.1.86

Gilinsky, A.S. (1951). Perceived size and distance in visual space. *Psychological Review, 58*(6), 460-482. https://doi.org/10.1037/h0061505

Glasauer, S., Amorim, M.-A., Viaud-Delmon, I., & Berthoz, A. (2002). Differential effects of labyrinthine dysfunction on distance and direction during blindfolded walking of a triangular path. *Experimental Brain Research, 145*, 489-497. https://doi.org/10.1007/s00221-002-1146-1

Gogel, W.C. (1965). Equidistance tendency and its consequence. *Psychological Bulletin, 64*(3), 153-163. https://doi.org/10.1037/h0022197

Gogel, W.C. (1974). Cognitive factors in spatial responses. *Psychologia, 17*(4), 213-225.

Gogel, W.C. (1977). An indirect measure of perceived distance from oculomotor cues. *Perception &*

Psychophysics, 21(1), 3-11. https://doi.org/10.3758/BF03199459

Gogel, W.C., & Da Silva, J.A. (1987). A two-process theory of the response to size and distance. *Perception & Psychophysics, 41*(3), 220-38. https://doi.org/10.3758/BF03208221

Gogel, W.C., & Tietz, J.D. (1973). Absolute motion parallax and the specific distance tendency. *Perception & Psychophysics, 13*(2), 284-292. https://doi.org/10.3758/BF03214141

Gomes, B.C., Oliveira, L.E.M.P., Matsushima, E.H., Santos, M.B., Ribeiro-Filho, N.P., & Da Silva, J.A. (1999). Produzindo distâncias para evitar colisão contra um obstáculo fixo em ambiente rígido? *Paidéia: Cadernos de Psicologia e Educação, 9*(17), 8-13. https://doi.org/10.1590/S0103-863X1999000200002

Goodale, M.A., & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20-25. https://doi.org/10.1016/0166-2236(92)90344-8

Künnapas, T. (1968). Distance perception as a function of available visual cues. *Journal of Experimental Psychology, 77*(4), 523-529.

Lackner, J.R., & DiZio, P. (2005). Vestibular, proprioceptive, and haptic contributions to spatial orientation. *Annual Review of Psychology, 56*, 115-147. https://doi.org/10.1146/annurev.psych.55.090902.142023

Leclere, N.X., Sarlegna, F.R., Coello, Y., & Bourdin, C. (2019). Sensori-motor adaptation to novel limb dynamics influences the representation of peripersonal space. *Neuropsychologia, 131*, 193-204. https://doi.org/10.1016/j.neuropsychologia.2019.05.005

Lederman, S.J., Klatzky, R.L., Collins, A., & Wardell, J. (1987). Exploring environments by hand or foot: a time-based heuristics for encoding distance in movement space. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 606-614. https://doi.org/10.1037//0278-7393.13.4.606

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

Lee, D.N., & Lishman, J.R. (1977a). Visual proprioceptive control of stance. *Journal of Human Movement Studies, 1*, 87-95.

Lee, D.N., & Lishman, J.R. (1977b). Visual control of locomotion. *Scandinavian Journal of Psychology, 18*,

224-230. https://doi.org/10.1111/j.1467-9450.1977.tb00281.x

Loomis, J.M., & Beall, A.C. (1998). Visually controlled locomotion: its dependence on optic flow, three-dimensional space perception, and cognition. *Ecological Psychology, 10*(3-4), 271-285. https://doi.org/10.1080/10407413.1998.9652685

Loomis, J.M., Da Silva, J.A., Fujita, N., & Fukusima, S.S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 906-921. https://doi.org/10.1037//0096-1523.18.4.906

Loomis, J.M., & Philbeck, J.W. (2008). Measuring spatial perception with spatial updating and action. In R.L. Klatzky, B. MacWhinney, & M. Behrmann (Eds.), *Embodiment, Ego-Space, and Action* (pp. 1-43). New York, NY: NY Psychology Press.

Manzone, J., & Heath, M. (2018). Goal-directed reaching: the allocentric coding of target location renders an offline mode of control. *Experimental Brain Research, 236*, 1149-1159. https://doi.org/10.1007/s00221-018-5205-7

Marken, R.S. (1986). Perceptual organization of behavior: a hierarchical control model of coordinated action. *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 267-276. https://doi.org/10.1037//0096-1523.12.3.267

Matsushima, E.H. (2004). Are perception and action responses really dissociated? Revisiting dorsal/ventral pathways hypothesis. In A.M. Oliveira, M.P. Teixeira, G.F. Borges & M.J. Ferro (Eds.), *Proceedings of the Twentieth Annual Meeting of the International Society for Psychophysics – Fechner Day 2004* (p. 204-209). Coimbra, Portugal: ISP.

Matsushima, E.H., Chiaretti, P., Kreling, D.B., Lima, M.F., Da Silva, J.A., & Ribeiro-Filho, N.P. (2004). Um invariante no controle da percepção e ação em tarefas de bissecção. *Paidéia: Cadernos de Psicologia e Educação, 14*(27), 83-88. https://doi.org/10.1590/S0103-863X2004000100011

Matsushima, E.H., Gomes, B.C., Ribeiro-Filho, N.P., & Da Silva, J.A. (2001). Do people walk through exocentric intervals or to perceived egocentric locations? In E. Sommerfeld, R. Kompass, & T. Lachmann

(Eds.), *Proceedings of the Seventeenth Annual Meeting of the International Society for Psychophysics* (pp. 523-528). Lengerich, Germany: Pabst Science Pubs/ISP.

Matsushima, E.H., & Ribeiro-Filho, N.P. (2003). Interações entre sistemas de referência alocêntricos e egocêntricos: evidências dos estudos com direção percebida. *Estudos e Pesquisas em Psicologia, 3*(1), 105-118.

Matsushima, E.H., Ribeiro-Filho, N.P., Douchkin, I.O., & Da Silva, J.A. (2002). Interaction between binocular and pictorial cues for visually directed walking. In J.A. Da Silva, E.H. Matsushima & N.P. Ribeiro-Filho (Eds.), *Proceedings of the Eighteenth Annual Meeting of the International Society for Psychophysics* (pp. 245-51). Rio de Janeiro, RJ: ISP.

Muroi, D., & Higuchi, T. (2017). Walking through an aperture with visual information obtained at a distance. *Experimental Brain Research, 235*, 219-230. https://doi.org/10.1007/s00221-016-4781-7

Olthuis, R., Van Der Kamp, J., & Caljouw, S. (2017). Verbalizations Affect Visuomotor Control in Hitting Objects to Distant Targets. *Frontiers in Psychology, 8*, 661. https://doi.org/10.3389/fpsyg.2017.00661

Paillard, J. (1991). Motor and representational framing of space. In J. Paillard (Ed.), *Brain and Space* (pp. 163-182). Oxford: Oxford University Press.

Philbeck, J.W., & Loomis, J.M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 72-85. https://doi.org/10.1037//0096-1523.23.1.72

Philbeck, J.W. Loomis, J.M., & Beall, A.C. (1997). Visually perceived location is an invariant in the control of action. *Perception & Psychophysics, 59*(4), 601-612. https://doi.org/10.3758/bf03211868

Pierce, J.E., Saj, A., & Vuilleumier, P. (2019). Differential parietal activations for spatial remapping and saccadic control in a visual memory task. *Neuropsychologia, 131*, 129-138. https://doi.org/10.1016/j.neuropsychologia.2019.05.010

Redding, G.M., & Wallace, B. (2001). Calibration and alignment are separable: evidence from prism adaptation. *Journal of Motor Behavior, 33*(4), 401-412. https://doi.org/10.1080/00222890109601923

Rieser, J.J., Pick, Jr., H.L., Ashmead, D.H., & Garing, A.E. (1995). Calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 480-497. https://doi.org/10.1037//0096-1523.21.3.480

Rieser, J.J., Ashmead, D.H., Talor, C.R., & Youngquist, G.A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception, 19*, 675-689. https://doi.org/10.1068/p190675

Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A. & Serences, J.T. (2018). Expectations do not alter early sensory processing during perceptual decision-making. *The Journal of Neuroscience, 38*(24), 5632-5648. https://doi.org/10.1523/JNEUROSCI.3638-17.2018

Schneider, W.X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition, 2*(2-3), 331-376. https://doi.org/10.1080/13506289508401737

Schwartz, M. (1999). Haptic perception of the distance walked when blindfolded. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 852-865. https://doi.org/10.1037//0096-1523.25.3.852

Similä, S.S., & McIntosh, R.D. (2015). Look where you're going! Perceptual attention constrains the online guidance of action. *Vision Research, 110B*, 179-189. https://doi.org/10.1016/j.visres.2014.06.002

Steenhuis, R.E., & Goodale, M.A. (1988). The effects of time and distance on accuracy of target-directed locomotion: does an accurate short-term memory for spatial location exist? *Journal of Motor Behavior, 20*(4), 399-415. https://doi.org/10.1080/00222895.1988.10735454

Sun, H.-J., Campos, J.L., Young, M., Chan, G.S.W., & Ellard, C.G. (2004). The contributions of static visual cues, nonvisual cues, and optic flow in distance estimation. *Perception, 33*, 49-65. https://doi.org/10.1068/p5145

Thomson, J.A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance, 9*(3), 427-443.

https://doi.org/10.1037//0096-1523.9.3.427

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549-586). Cambridge, MA: MIT Press.

Vishton, P.M., Rea, J.G., Cutting, J.E., & Nuñez, L.N. (1999). Comparing effects of the horizontal-vertical illusion on grip scaling and judgment: relative versus absolute, not perception versus action. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1659-1672. https://doi.org/10.1037//0096-1523.25.6.1659

Weiss, P.H., Marshall, J.C., Zilles, K., & Fink, G.R. (2003). Are action and perception in near and far space additive or interactive factors? *NeuroImage, 18*, 837-846. https://doi.org/10.1016/s1053-8119(03)00018-1

Zhao, H., & Warren, W.H. (2015). On-line and model-based approaches to the visual control of action. *Vision Research, 110*, 190-202. https://doi.org/10.1016/j.visres.2014.10.008