



Late vision: processes and epistemic status

Athanassios Raftopoulos*

Department of Psychology, University of Cyprus, Nicosia, Cyprus

Edited by:

Arnon Cahen, Ben Gurion University
in the Negev, Israel
Michela C. Tacca, Heinrich-Heine
University Düsseldorf, Germany

Reviewed by:

Bence Nanay, University of Antwerp,
Belgium
Sebastian Watzl, Harvard University,
USA
Jack Lyons, University of Arkansas,
USA

***Correspondence:**

Athanassios Raftopoulos, Department
of Psychology, University of Cyprus,
P.O. Box 20537, Nicosia, Cyprus.
e-mail: raftop@ucy.ac.cy

In this paper, I examine the processes that occur in late vision and address the problem of whether late vision should be construed as a properly speaking perceptual stage, or as a thought-like discursive stage. Specifically, I argue that late vision, its (partly) conceptual nature notwithstanding, neither is constituted by nor does it implicate what I call pure thoughts, that is, propositional structures that are formed in the cognitive areas of the brain through, and participate in, discursive reasoning and inferences. At the same time, the output of late vision, namely an explicit belief concerning the identity and category membership of an object (that is, a recognitional belief) or its features, eventually enters into discursive reasoning. Using Jackendoff's distinction between visual awareness, which characterizes perception, and visual understanding, which characterizes pure thought, I claim that the contents of late vision belong to visual awareness and not to visual understanding and that although late vision implicates beliefs, either implicit or explicit, these beliefs are hybrid visual/conceptual constructs and not pure thoughts. Distinguishing between these hybrid representations and pure thoughts and delineating the nature of the representations of late vision lays the ground for examining, among other things, the process of conceptualization that occurs in visual processing and the way concepts modulate perceptual content affecting either its representational or phenomenal character. I also do not discuss the epistemological relations between the representations of late vision and the perceptual judgments they "support" or "guide" or "render possible" or "evidence" or "entitle." However, the specification of the epistemology of late vision lays the ground for attacking that problem as well.

Keywords: late vision, visual awareness, visual understanding, conceptualization, perceptual beliefs, essential indexicals

INTRODUCTION

In earlier work (Raftopoulos, 2009), I analyzed early vision, which I claimed is a pre-attentional visual stage unaffected by top-down conceptual/cognitive modulation. (In what follows when I refer to top-down processes I mean cognitively driven processes, although there is top-down flow of signals within the visual areas. In addition, where I refer to attention I mean cognitively driven attention, unless I state otherwise.) Thus, early vision is a cognitively impenetrable stage of visual processing. I have related the content of the states of early vision with the non-conceptual content (NCC) of perception by arguing that the cognitive impenetrability of some states and contents is a necessary and sufficient condition for these states and contents to be non-conceptual. I also underlined Pylyshyn's (2003) distinction between early vision and late vision. The latter is cognitively penetrated and involves the modulation of processing by either spatial or object/feature centered attention.

In this paper, I examine the processes that occur in late vision and discuss whether late vision should be construed as a perceptual stage or as a thought-like discursive stage. I argue that late vision, its (partly) conceptual nature notwithstanding, does not consist in pure thoughts, that is, propositional structures that are formed in the cognitive areas of the brain and participate in discursive reasoning and inferences. The content of the output of late vision, that

is, an explicit belief concerning the identity of an object (recognitional belief), enters into discursive reasoning. Using Jackendoff's (1989) distinction between visual awareness, which characterizes perception, and visual understanding, which characterizes pure thought, I claim that the contents of late vision belong to visual awareness and not to visual understanding. Although late vision implicates beliefs, either implicit or explicit, these beliefs are hybrid visual/conceptual constructs and not pure thoughts. Distinguishing between these hybrid representations and pure thoughts lays the ground for examining the conceptualization of perceptual content and the way concepts modulate it affecting either its representational or its phenomenal character. I do not discuss these problems here, as I do not discuss the epistemological relations between the representations of late vision and the perceptual judgments they "support" or "evidence" or "entitle." However, the specification of the epistemic status of late vision lays the ground for attacking this problem as well.

In the first section, I sketch early vision. Then, I discuss late vision with an emphasis on its role in object recognition. The purpose is to examine some of the contents and processes of late vision and their timing. In the third section, I argue that late vision should be considered as a perceptual rather than as a discursive stage involving understanding, that is, a stage of thought processing involving pure thoughts and inferences from

propositionally structured premises to the identity of objects. My argument is based on considerations regarding the sorts of contents and processes formed in early and late vision.

EARLY VISION

Early vision includes a feed forward sweep (FFS) in which signals are transmitted bottom-up. In visual areas (from LGN to IT) FFS lasts for about 100 ms. It also includes a stage at which lateral and recurrent processes that are restricted within the visual areas and do not involve signals from cognitive centers occur. Recurrent processing starts at about 80–100 ms and culminates at about 120–150 ms. Lamme (2003) calls it local recurrent processing (LRP). The unconscious FFS extracts high-level information that could lead to categorization, and results in some initial feature detection. LRP produces further binding and segregation. The representations formed at this stage are restricted to information about spatio-temporal and surface properties, color, texture, orientation, motion, and affordances of objects, in addition to the representations of objects as bounded, solid entities that persist in space and time.

By not involving signals from the cognitive areas of the brain, FFS and LRP are cognitively impenetrable/conceptually encapsulated, since the transmitting of signals within the visual system is not affected by top-down signals produced in cognitive areas. Early vision processing is not affected directly by top-down signals from cognitive states through attention – that is, attention does not affect the early visual processes although it may affect pre-perceptual and post-perceptual stages of vision. I have argued that this leads to the thesis that early vision has NCC, provided that concepts do not figure inherently in the perceptual system, a possibility that I have rejected (Raftopoulos, 2009). The processes during early vision that result in states with personal-level NCC correspond to Dretske's (1995) phenomenal seeing¹.

LATE VISION

The conceptually² modulated stage of visual processing is called late vision. Starting at 150–200 ms, signals from higher executive centers including mnemonic circuits intervene and modulate perceptual processing in the visual cortex and this signals the onset of global recurrent processing (GRP). In 50 ms low spatial frequency (LSF) information reaches the IT and in 100 ms high spatial frequency (HSF) information reaches the same area (Kihara and Takeda, 2010). (LSF signals precede HSF signals. LSF information is transmitted through fast magnocellular pathways, while HSF information is transmitted through slower parvocellular pathways.) Within 130 ms post-stimulus, parietal areas in the dorsal system but also areas in the ventral pathway (IT cortex) semantically process the LSF information and determine the gist of the scene based on stored knowledge that generates predictions about

the most likely interpretation of the input, even in the absence of focal attention.

This information reenters the extrastriate visual areas and modulates (at about 150 ms) perceptual processing facilitating the analysis of HSF, for example, by specifying certain cues in the image that might facilitate target identification (Barr, 2009; Kihara and Takeda, 2010; Peyrin et al., 2010). Determining the gist may speed up the FFS of HSF by allowing faster processing of the pertinent cues, using top-down connections to preset neurons coding these cues at various levels of the visual pathway (Delmore et al., 2004). Thus, at about 150 ms, specific hypotheses regarding the identity of the object(s) in the scene are formed using HSF information in the visual brain and information from visual working memory (WM). The hypothesis is tested against the detailed iconic information stored in early visual circuits including V1. ERP's waveforms that distinguish scenes and objects in object recognition tasks are registered at about 150 ms in extrastriate areas and are thought to be early indices of P3³ (Fabre-Thorpe et al., 2001; Johnson and Olshausen, 2005). This testing requires that top-down signals reenter the early visual areas of the brain, and mainly V1. Indeed, evidence shows that V1 is reentered by signals from higher cognitive centered mediated by the effects of object/feature centered attention at 235 ms post-stimulus (Chelazzi et al., 1993; Roelfsema et al., 1998). This leads to the recognition of the object(s) in the visual scene. This occurs, as signaled by the P3 ERP waveform, at about 300 ms in the IT cortex, whose neurons contribute to the integration of LSF and HSF information.

A detailed analysis of the form that the hypothesis testing might take is provided by Kosslyn (1994). Note that one need not subscribe to some of the assumptions presupposed by Kosslyn's account (see Raftopoulos, 2010 for criticism), but these disagreements do not undermine the framework. Suppose that one sees an object. A retinotopic image is formed in the visual buffer, which is a set of visual areas in the occipital lobe that is organized retinotopically. An attentional window selects the input from a contiguous set of points for detailed processing. This is allowed by the spatial organization of the visual buffer. The information included in the attention window is sent to the dorsal and ventral system where different features of the image are processed. The ventral system retrieves the features of the object, whereas the dorsal system retrieves information about the location, orientation, and size of the object. Eventually, the shape, the color, and the texture of the object are registered in anterior portions of the ventral pathway. This information is transmitted to the pattern activation subsystems in the IT cortex where the image is matched against representations stored there, and the compressed image representation of the object is thereby activated. This representation (which is an hypothesis regarding the identity of an object) provides imagery feedback to the visual buffer where it is matched against the input image to test the hypothesis against the fine pictorial details registered in the retinotopical areas of the visual buffer. If the match is satisfactory, the category pattern activation subsystem sends the relevant pattern code to associative

¹In Raftopoulos (2009) I argue that a state with NCC does not have a propositional content, and that two states cannot have the same content and one have NCC and the other conceptual content. In this paper, I assume both theses. I also assume that part of the NCC is content at the personal level and that one has *phenomenal awareness* of that content.

²Concepts are constant, context independent, and freely repeatable elements that figure constitutively in propositional contents; they correspond to lexical items.

³The P3 waveform is elicited about 250–600 ms and is generated in many areas in the brain and is associated with cognitive processing and the subjects' reports. P3 may signify the consolidation of the representation of the object(s) in working memory.

or WM, where the object is tentatively identified with the help of information arriving at the WM through the dorsal system (information about, size, location, and orientation). Occasionally the match in the pattern activation subsystems is enough to select the appropriate representation in WM. On other occasions, the input to the ventral system does not match well a visual memory in the pattern activation subsystems. Then, a hypothesis is formed in WM. This hypothesis is tested with the help of other subsystems (including cognitive ones) that access representations of such objects and highlight their more distinctive feature. The information gathered shifts attention to a location in the image where an informative characteristic can be found. The attention window zooms on object's distinctive feature, and the pattern code for it is sent to the pattern activation subsystem and to the visual buffer where a second cycle of matching commences.

ERP experiments registering the time onset of various waveforms related to specific processes in the brain largely confirm this analysis. The N2 ERP component that signifies cognitively driven spatial-attentional effects on the extrastriate cortex is registered at about 170–200 ms. Thus, by 170 ms spatial attention directly modulates visual processing. However, cognitive top-down modulation of the extrastriate cortex, mainly V4, from the IT and parietal cortex is found as early as 150 ms, which, as we saw, is the first sign of the process of object identification.

Eventually there is considerable competition since only few items can enter in interactions with the higher hierarchically processing levels. Further selection becomes necessary when several stimuli reach the brain but only one response is possible. Attentional selection intervenes to resolve this competition. The selection results from the combination of bottom-up information processing with WM and long-term memory (LTM) that recover the meaning of input and relate it to the subject's current goals. In the biased competition account of attention (Desimone and Duncan, 1995), attention is the competition between neuronal populations that encode environmental stimuli. All the stimuli in a visual scene are initially processed in parallel and activate neuronal assemblies that represent them. These assemblies eventually engage in competitive interactions for several reasons (when, for example, some behaviorally relevant feature or object must be selected among all present stimuli).

Recurrent interactions with areas outside the visual stream make storage in visual WM possible and give rise to GRP. In GRP, standing knowledge, that is, information stored in the synaptic weights of the neurons is activated (becoming part of WM) and modulates visual processing, which up to that point was conceptually encapsulated. Consequently, during GRP the conceptualization of perceptual content starts and the states formed during this stage have (perhaps partly) conceptual and eventually propositional contents⁴. This is the stage where the 3D sketch is formed,

⁴This means that some conceptual content in late vision may not be propositionally structured, although recognitional beliefs have propositional structure. It is also possible that some states in late vision have both NCC and conceptual content. I will not elaborate on these issues here. Note that if some of the states of late vision can have conceptual contents that are not propositionally structured, my thesis that late vision does not involve inferences is strengthened because inferences relate propositional structures.

since the recovery of the 3D sketch, that is, the representation of an object independently of the viewer's perspective, cannot be the output of early vision. This recovery cannot be purely data-driven, since what is regarded as an object depends on the subsequent usage of the information, and thus depends on the knowledge about objects. It follows that the formation of the 3D sketch requires constitutively the application of concepts⁵. Seeing 3D sketches of objects is an instance of amodal perception, i.e., the representation of object parts or features that are not visible from the viewer's standpoint. Thus, late vision involves a synergy of perceptual bottom-up processing and top-down processing, where knowledge from past experiences guides the formation of hypotheses about the identity of objects present in the visual scene. Late vision is also responsible for the experience of the 3D sketch.

There are two sorts of completion. In modal completion the viewer has a distinct visual impression of a hidden contour or other hidden features even though these features are not occurring sensory features. The perceptual system fills in the missing features, which thus become as phenomenally occurring as the occurring sensory features of the object. In amodal completion, one does not have a perceptual (imagination is not perception) impression of the object's hidden features since the perceptual system does not fill in the missing features as it happens in modal perception; the hidden features are not perceptually occurring.

There are cases of amodal perception that are purely perceptual, that is, bottom-up. In these cases, although no direct signals from the hidden features impinge on the retina (there is no local information available), the perceptual system can extract information regarding them from the global information contained in the visual scene without any cognitive involvement, as the resistance of the ensuing percepts to beliefs indicates. However, in such cases, the hidden features are not perceived. One simply has the visual impression of a single concrete object that is partially occluded and not the visual impression of various disparate image regions. Therefore, in these perceptually driven amodal completions there is no mental imagery involved, since no top-down signals from cognitive areas are required for the completion, and since the hidden features are not phenomenologically present.

There are also cases of amodal completion that are cognitively driven (Briscoe, 2011 calls them C-completions⁶), such as the formation of the 3D sketch of an object, in which the hidden features of the object are represented through the top-down activation of the visual cortex from the cognitive centers of the brain. In some of these cases, top-down processes activate the early visual areas and fill in the missing features that become phenomenologically present. In other cases of C-completion, the viewer simply forms a pure thought concerning the hidden structure in the absence of any activation of the visual areas and, thus, in the absence of

⁵The view that the formation of the viewer independent representation of an object relies on object knowledge is common in theories of the formation of the 3D viewer independent representation. Biederman (1987) thinks that object recognition is based on part decomposition, which is the first stage in forming a structural description of an object. This decomposition cannot be determined by general principles reflecting the structure of the world alone, since the decomposition appears to depend upon knowledge of specific objects.

⁶Briscoe's paper analyzes Nanay's (2010) account of the role of imagination in amodal completion.

mental imagery. As the latter possibility may threaten my thesis that C-completion takes place in late vision, I discuss it in the next section.

Before I proceed, allow me to delve on “mental imagery,” since the way it is used may cause some confusion concerning the top-down processes in late vision. Imagery is central in Kosslyn’s (1994) account of object recognition. As we saw, Kosslyn thinks that visual imagery is involved in all cases of perception and covers all the top-down flow of information either from the associative areas of the brain or the pattern activation subsystems in the IT cortex. Strawson (1974) also holds that object recognition involves visual imagery. Discussions on amodal completion emphasize the role of imagery in completing the hidden features by representing them and occasionally making them phenomenologically present even though they are perceptually absent (Nanay, 2010)⁷. In discussing late vision, I emphasized the role of top-down processes that are necessary for object recognition. Now, it is well known that many of the neural systems engaged in mental imagery are also actively involved in the formation of the percept, most notably the early visual areas. Since mental imagery is usually related to top-down processes, imagery could be assimilated to late vision, which involves top-down processes too. As mental imagery involves top-down activation of the visual areas, it is tempting to claim that the top-down processes in late vision are instances of visual imagery, especially so in the case of C-completion in which the object or feature that is represented through mental imagery is absent from the visual scene.

To decide the issue one should define mental imagery. Usually mental imagery is related to the mental construction of the image of an object or feature in its absence. The image formed from actual (perceptual) experience is called a percept to distinguish this image from an imagined or mental image. When a subject is asked to recall a visual object, the image formed in memory is called a mental image. The mental image is constructed via top-down processes (when, for example, subjects are presented with a lower case letter and are asked to form a mental image of the upper case letter, a task that is cognitively driven since it requires knowledge of the upper case letter), while the percept is constructed through a synergy of top-down and bottom-up processes. Thus, mental imagery is usually construed as (i) involving only top-down cognitively driven processes, and (ii) taking place in the absence of the imagined object or feature. This is how I use the term.

Kosslyn (1994) and Strawson (1974), in contrast, uses the term to designate the top-down processes in object recognition. Kosslyn talks about imagery feedback to the visual buffer both from the associative concept involving areas of the brain, and the pattern activation subsystems that Kosslyn thinks store non-conceptualized information. Therefore, mental imagery can be either cognitively driven or data-driven, which goes against the usual construal of mental imagery. Moreover, mental imagery is engaged in perceptual tasks of object recognition, which means that Kosslyn foregoes the second trait of mental imagery as

well. Nanay (2010, pp. 244–246, 250) uses visual imagery to account for cognitively driven amodal completion, and specifically, to designate the top-down knowledge-driven effects on visual processing.

Mental imagery is perceptually and not propositionally coded, even though it may start with the activation of concepts in associative memory (Kosslyn, 1994). However, the activation of the visual areas in a top-down manner in mental imagery is not the same as the activation of these same areas by sensory signal. For example, the top-down induced activation in the absence of retinal input is weaker and, thus, the modal “mode” associated with mental imagery is not as strong or lively as in perception. Although it is true that when an object is imagined as opposed to merely thought about a number of properties must be added to the description, these properties fall far short of all those that would be present in perception. Not only some features may be omitted, but also precise iconic and metric information is lost in mental imagery. Since the concepts that activate the visual cortex represent abstract categorical information, such as bright, red, and not the determinate color say red21 (which is why one cannot recall the determinate color of an object but only its category membership), not all visual details of the actual visual scene can be the contents of a state of visual imagery (Raftopoulos, 2010). In late vision, on the other hand, the presence of the visual object allows conceptual demonstratives to rely on the presence of the sample and overcome any conceptual limitations.

Since late vision constitutively involves a synergy of bottom-up and top-down processing, whereas mental imagery, as I construe it, involves only top-down flow of information to early visual areas in the absence of sensory stimulation, I prefer (pace Kosslyn and Nanay) not to use “imagery” to designate the top-down activation of the visual cortex in late vision, even in those cases in which top-down processing completes hidden features of objects. Mental imagery differs from seeing in that it uses only the late processing components of the perceptual system when the early processing sensory-driven processes are unavailable (as when there is no sensory stimulation). Visual imagery activates the (inactive) visual processing areas to recreate to a certain extent a visual scene. As such, mental imagery, unlike late vision, involves only top-down processes. Although in both cases the early visual areas are reentered from signals emanating from cognitive centers, in late vision the cognitive centers are activated through bottom-up signals from the visual cortex, while in visual imagery the cognitive centers are activated in the absence of any sensory stimulation on the retina. Thus, I think that the top-down processes in late vision should be distinguished from mental imagery in that the former are essentially engaged by the existence of sensory stimuli on the retina, whereas in the latter there are no sensory stimuli.

IS LATE VISION A VISUAL STAGE OR A DISCURSIVE THOUGHT-LIKE STAGE?

THE PROBLEM

Jackendoff (1989) distinguishes visual awareness from visual understanding. There is a qualitative difference between the experience of a 3D sketch and the experience of a 2D/2D sketch. One is aware of the 3D sketch or of category based representations,

⁷The phenomenal/non-phenomenal distinction is orthogonal to the discussion on mental imagery since mental imagery, exactly like perception, can either be accompanied by consciousness, or it can be implicit (as in implicit perception). I wish to thank a reviewer for suggesting this.

however, this is not visual awareness but some other kind of awareness. Visual awareness is awareness of Marr's 21/2D sketch, which is the viewer-centered representation of the visible surfaces of objects, while the awareness of the 3D sketch is visual understanding. Thus, the 3D sketch, which includes the unseen surfaces that are not represented in the 21/2D sketch, is a result of an inference; amodal completion is an inference. Jackendoff's views belong to the so-called belief-based account of amodal completion: the 3D sketch is the result of beliefs inferred from the object's visible features and other background information from past experiences.

The problem is whether object identification and C-completion that occur in late vision and are both dependent on concepts should be thought of as cases of vision or as cases of discursive understanding involving inferences. If late vision involves conceptual contents and if the role of concepts and stored knowledge consists, among other things, in providing some initial interpretation of the visual scene and in forming hypotheses about the identity of objects that are tested against perceptual information, one is tempted to say that this stage relies on inferences (this is what hypothesis testing amounts to) and, thus, differs in essence from the purely perceptual processes of early vision. Perhaps it would be better to construe late vision as a discursive stage involving thoughts, in the way of Jackson's (1977) epistemic seeing, where "seeing" is used in a metaphorical non-perceptual sense, as where one says of his friend whom she visited "I see he has left," based on perceptual evidence. It is, also possible that Dretske (1993, 1995) thinks that seeing in the doxastic sense is not a visual but, rather, a discursive stage.

One might object that abandoning this usage of "to see" violates ordinary usage. A fundamental ingredient of visual experience consists in meaningful 3D solid objects. Adopting this proposal would mean that one should resist talking of seeing tigers and start talking about seeing viewer-centered visible surfaces. "By this criterion, much of the information we normally take to be visually conscious would not be, including the 3D shape of objects as well as their categorical identity" (Palmer, 1999, p. 649).

The arguments to common language notwithstanding, I think that one should not assume either that late vision is an inferential discursive stage that constitutively involves thoughts in the capacity of premises in inferences whose conclusion is the content of the states of late vision (although implicit hypotheses play a role), or that late vision consists in discursively entertaining thoughts. The reason is twofold. First, I think that seeing an object is not the result of an inference, that is, a movement in thought from some premises to a conclusion and, thus, a discursive process, even though it involves concepts. Second, late vision is a stage in which conceptual modulation and perceptual processes form an inextricable link that differentiates late vision from discursive stages and renders it a different sort of a set of processes than understanding, even though late vision involves implicit beliefs regarding objects that guide the formation of hypotheses concerning object identity, and an explicit belief of the form "that O is F" eventually arises in the final stages of late vision. Late vision has an irreducible visual ingredient, which makes it different from discursive understanding. Before I discuss all these, let me clarify some terminological issues.

Beliefs

Traditionally judgments are occurrent states, whereas beliefs are dispositional states. To judge that O is F is to predicate Fness to O, while endorsing the predication (McDowell, 1994). To believe that O is F is to be disposed to judge, under the right circumstances, that O is F. This is the first sense in which beliefs are dispositional items. Now, as the reader recalls, I have distinguished between standing knowledge – information stored in LTM – and information that is activated in WM. The belief that O is F may be a standing information in LTM, a memory, because, say, one has seen O to be F in the past, even though presently one does not have an occurrent thought about O. Beliefs need not be consciously or unconsciously recalled or apprehended in order to be possessed by a subject, which means that beliefs are dispositional rather than occurrent items; this is a second sense in which beliefs are dispositional. When this information is activated, the occurrent thought that O is F emerges in WM. In the literature one finds the distinction between "thought" and "standing knowledge" (Prinz, 2002, p. 148). Accordingly, all thoughts are occurrent states by being activated in WM. Thus, I use "occurrent thought" and "thought" as synonymous.

It follows that a belief *qua* dispositional state may be either a piece of standing knowledge, in which case it is dispositional in the sense that when activated it becomes a thought, or a thought that awaits endorsement to become a judgment, in which case the belief is dispositional in the sense that it has the capacity to become a judgment. In the first case, if beliefs are stored in LTM as standing knowledge and if thoughts are occurrent states, beliefs are not the same as thoughts although a belief when activated becomes a thought. In the second case, a belief is a thought held in WM, albeit one that has not been yet endorsed. There are interesting epistemological implications but they are irrelevant here. In what follows, I assume that beliefs are either thoughts or pieces of standing information, which have not been endorsed and, thus, are not judgments. One might wonder how is it possible to understand a belief as an occurrent thought that is not endorsed? An explanation has to wait until I have explained why late vision does not involve inferences.

State consciousness

It is important for the discussion that follows to clarify another problem, namely, under which conditions are beliefs conscious or not. An intuitive answer is that, as a matter of course, one may entertain beliefs or judgments and use them for various purposes (for example to draw conclusions in inferences or guide actions) even though one is not conscious that one entertains these beliefs or judgments (as in the case of using implicit premises in an argument); these beliefs are implicit. Underneath this intuitive view one discerns the assumption that a state is conscious if the person who has it is conscious that she is in that state. Either that person has a second order thought that she is entertaining such a belief – that is, she has fact-awareness that she is entertaining that state – or she has a second order experience or inner sense that she is in such a state – that is, she has thing-awareness of the state – where "thing-awareness" and "fact-awareness" are used in the way Dretske (1993) defines them. If one subscribes to this view, what makes a mental state of a person conscious is the person's

awareness of the state. However, Dretske (1993) argues that what renders a person's state conscious is not some sort of second order awareness that one is in such and such state, or that she is having that state. A state is conscious

being a certain sort of representation, it makes one aware of the properties (of *x*) and objects (*x* itself) of which it is a sensory representation. . . [A] certain belief is conscious, not because the believer is conscious of it (or conscious of having it), but because it is a representation that makes one conscious of the fact (that *P*) that it is a belief about. . . beliefs are conscious, not because you are conscious of them, but because, so to speak, you are conscious *with* them (Dretske, 1993, pp. 437–438).

Beliefs that are thought of as implicit but play a cognitive role in making a person aware of some facts or things are conscious (a first-order consciousness). Dretske does not claim that everything that happens to one when one becomes conscious of some object or event is conscious. However, a perceptual experience or a belief has to be conscious in order for a person to be made aware of things and events. I do not assess Dretske's thesis, which is only among many views on consciousness (some of which are higher-order theories that Dretske resists), and I remain neutral as to how conscious state should be construed. By "implicit belief" I mean the belief held by a person who is not aware that she is having that belief.

Inference

My claim is that the processes in late vision are not inferential processes where "inference" is understood as discursive, that is, as a process that involves drawing propositions—conclusions from other propositions acting as premises by applying (explicitly or implicitly) inferential rules that are also represented. These inferences are distinguished from "inferences" as understood by vision scientists according to whom any transformation of signals carrying information according to some rule is a form of inference. "Every system that makes an estimate about unobserved variables based on observed variables performs inference. . . We refer to such inference problems that involve choosing between distinct and mutually exclusive causal structures as causal inference" (Shams and Beierholm, 2010).

LATE VISION, HYPOTHESIS TESTING, AND INFERENCE

I think that the states of late vision are not inferences from premises that include the contents of early vision states, even though it is usual to find claims that one infers that a tiger, for example, is present from the perceptual information retrieved from a visual scene. An inference relates some propositions in the form of premises with some other proposition, the conclusion. However, the objects and properties as they are represented in early vision do not constitute contents in the form of propositions, since they are part of the non-propositional NCC of perception. In late vision, the perceptual content is conceptualized but the conceptualization is not a kind of inference but rather the application of stored concepts to some input that enters the cognitive centers of the brain and activates concepts by matching their content. Thus, even though the states in late vision are

formed through the synergy of bottom-up visual information and top-down conceptual influences, they are not inferences from perceptual content.

Late vision involves hypotheses regarding the identity of objects and their testing against the sensory information stored in iconic memory. One might think that inferences are involved since testing hypotheses is an inferential process even though it is not an inference from perceptual content to a recognitional thought. It is, rather, an argument of the form if *A* and *B* then (conclusion) *C*, where *A* and *B* are background assumptions and the hypothesis regarding the identity of an object respectively, and *C* is the set of visual features that the object is likely to have. *A* consists of implicit beliefs about the features of the hypothesized visual object. If *C* is what obtains in the visual areas, that is, if the predicted visual features match those that are stored in iconic memory then the hypothesis about the identity of the object is likely correct. However, the test basis or evidence against which these hypotheses are tested for a match, that is, the iconic information stored in the sensory visual areas, is not a set of propositions but patterns of neuronal activations whose content is non-propositional.

There is nothing inference-like in this matching. It is just a comparison between the activations of neuronal assemblies that encode the visual features in the scene and the activations of the neuronal assemblies that are activated top-down from the hypotheses. If the same assemblies are activated then there is a match. If they are not, the hypothesis fails to pass the test. This can be done through purely associational processes of the sort employed, say, in connectionist networks that process information according to rules and, thus, can be thought of as instantiating processing rules, without either representing these rules or operating on language-like symbolic representations. Since inferences are carried out through rules that are represented in the system, and operate on symbolic structures, the processing in a connectionist network does not involve inferences, although it can be described in terms of inference making. Thus, even though seeing an object in late vision involves the application of concepts that unify the appearances of the object and of its features under some category, it is not an inferential process. The processes in late vision despite their reliance on background beliefs do not entail by themselves a recognitional belief.

Spelke (1988, p. 458)⁸ argues that "perceiving objects may be more akin to thinking about the physical world than to sensing the immediate environment." The reason is that the perceptual system, to solve the underdetermination problem of both the distal object from the retinal image and of the percept from the retinal image, employs a set of object principles and that reflect the geometry and the physics of our environment. Since the contents of these principles consist of concepts, and thus, the principles can be thought of as some form of knowledge about the world, perception engages in discursive, inferential processes. Against this, I have argued (Raftopoulos, 2009) that for a variety of reasons the processes that constrain the operations of the visual system should not be construed as inferences. Rather, they constitute the

⁸Spelke echoes Rock's (1983) views that the perceptual system combines inferentially information to form the percept. For example, from visual angle and distance information, one infers and perceives size.

modus operandi of the perceptual system, they are hardwired in the perceptual circuits, and they are not represented anywhere.

Being hardwired is another reason why perceptual processes should not be assimilated to inference making. Inferences presuppose that the subject applies explicitly or implicitly inferential rules that are represented in the subject. But the operations by means of which signals are transformed from one into the other in the visual system are not represented at all; they are just hardwired in the perceptual system. For this reason, perceptual operations should not be construed as inference rules, although they are describable in terms of inference rules.

LATE VISION AND DISCURSIVE UNDERSTANDING

Even if I am right that seeing in late vision is not the result of a discursive inference, it is still arguable that late vision should be better construed as a stage of discursive understanding rather than as a visual stage. If object recognition involves forming a belief about class-membership, even if the belief is not the result of an inference, why not say that recognizing an object is an experience-based belief that is a case of understanding rather than vision.

Late vision is more than object recognition

A first problem with this view is that late vision involves more than a recognitional belief. Suppose that S sees an animal and recognizes it as a tiger. In the parallel preattentive early vision, the proto-object that corresponds to the tiger is being represented amongst the other objects in the scene. The relevant activations enter the parietal and temporal lobes, and the prefrontal cortex, where the neuronal assemblies encoding the information about tigers are activated and this activation spreads through top-down signals to the visual areas of the brain where visual sensory memory stores the proto-objects extracted from the visual scene. The cells encoding the proto-object corresponding to the animal and its properties have their activations strengthened and win the competition against the assemblies encoding the proto-objects corresponding to the other objects in the scene. After a proto-object has been selected, the object recognition system forms hypotheses regarding the identity of the object. However, for the subject's confidence to reach the threshold that will allow her to form beliefs about the identity of the object and report it, these hypotheses must be tested (Treisman, 2006).

To test these hypotheses the visual system allocates resources to features and regions that would confirm or disconfirm the hypotheses. Conceptual information about a tiger affects visual processing and after some hypothesis testing the animal is recognized as a tiger through the synergy of visual circuits and WM. At this point the explicit belief "O is F" is formed. This occurs after 300 ms, when the viewer consolidates the object in WM and identifies it with enough confidence to report it, which means that beliefs are formed at the final phases of late vision. However, semantic modulation of visual processing and the process of conceptualization that eventually leads to object recognition starts at about 130–200 ms. There is, thus, a time gap, between the onset of conceptualization and the recognition of an object, which is a prerequisite for the formation of an explicit recognitional belief.

As Treisman and Kanwisher (1998) observe, although the formation of hypotheses regarding the categorization of objects

can occur within 130–200 ms after stimulus onset (the time depends on the saliency of the object), it takes another 100 ms for subsequent processes to bring this information into awareness so that the perceiver could be aware of the presence of an object and be able to report it. To form the recognitional belief that "O is F" one must be aware of the presence of an object token and construct first a coherent representation. This requires the enhancement through attentional modulation of the visual responses in early visual circuits that encode rich sensory information in order to integrate them into a coherent representation, which is why beliefs are delayed in time compared with the onset of conceptualization. It follows that not all of late vision involves explicit beliefs.

Late vision and consciousness

The beliefs involved in late vision in the form of hypotheses are not in the stream of the perceiver's consciousness; they are not explicit. The processing in late vision is done automatically and is outside both of the cognitive control of the viewer and of her awareness. Matching an input to a stored template is not under anyone's cognitive control and is not a process of which one is aware; neither is the determination of the gist of a visual scene. The conceptualization of the content of perception is not under anyone's control. Furthermore, for a thought to be conscious the person who has it must have access awareness to the contents of the thought; the perceiver reports, as it were, the content of her thoughts to herself. Thus, she must have some kind of a higher-order thought about the contents of her thought. Such a higher-order thought is not required in order to be able to recognize objects. Report awareness occurs in 500 ms, when the object has been categorized. This marks a difference between late vision and thought. Most of the contents and their transformations that occur during late vision cannot be in the realm of awareness, although the outcome of late vision is. Propositional inferences, by contrast, can be available to awareness.

Late vision as a synergy of bottom-up and top-down information processing

A third reason why the beliefs formed in late vision are partly visual constructs and not pure thoughts is that the late stage of late vision in which explicit beliefs concerning object identity are formed constitutively involves visual circuits (that is, brain areas from LGN to IT in the ventral system). Pure thought involves primarily an amodal form of representation formed in higher centers of the brain, even though these amodal representations can trigger in a top-down manner the formation of mental images and can be triggered by sensory stimulation. The point is that amodal representations can be activated without a concomitant activation of the visual cortex (see Prinz, 2002 notion of default concepts that are amodal representations). Perceptual representations in late vision, in contrast, are modal since they constitutively involve visual areas. Thus, what distinguishes late vision beliefs and pure thoughts is not so much their modal or amodal character (pure thoughts can also be accompanied by some sort of phenomenology), as the fact that the beliefs in late vision are formed through a synergy of bottom-up and top-down activation and their maintenance requires the active participation of the visual circuits. Pure

thoughts can be activated and maintained in the absence of any activation in visual circuits.

The constitutive reliance of late vision on the visual circuits suggests that late vision relies on the presence of the object of perception; it cannot cease to function as a perceptual demonstrative that refers to the object of perception, as this has been individuated though the processes of early vision (Raftopoulos and Muller, 2006; Burge, 2010, p. 542). As such, late vision is constitutively context dependent since the demonstration of the perceptual particular is always context dependent. Thought, on the other hand, by its use of context independent symbols, is free of the particular perceptual context. Even though both recognitional beliefs in late vision and pure perceptual beliefs involve concepts (pure attributive elements Burge, 2010), the concepts function differently in the two contexts. As Burge (2010, p. 545) claims “perceptual belief makes use of the singular and attributive elements in perception. In perceptual belief, pure attribution is separated from, and supplements, attributive guidance of contextually purported reference to particulars. . . Correct conceptualization of a perceptual attributive involves taking over the perceptual attributive’s range of applicability and making use of its (perceptual) mode of presentation.”

Note that the attributive and singular elements in perception correspond to the perceived objects and their properties and not to concepts concerning these objects and properties. The attributive elements (properties in perception) guide the contextual reference to particulars (the objects of perception) since the referent in a demonstrative perceptual reference is fixed through the properties of the referent as these properties are presented in perception – what I have called the non-conceptual mode of presentation of the object in perception (Raftopoulos and Muller, 2006). As such, they belong to the NCC content of perception (Burge, 2010, p. 538) Concepts enter the game in their capacity as pure attributions that make use of the perceptual mode of presentation. Burge’s claim that in perceptual beliefs pure attributions supplement attributions that are used for contextual reference to particulars may be read to mean that perceptual beliefs are hybrid states involving both visual elements (the contextual attributions used for determining reference to objects and their properties) and conceptualizations of these perceptual attributives in the form of pure attributions. In this case, the role of perceptual attributives is ineliminable and, thus, Burge’s perceptual beliefs map onto my recognitional beliefs of late vision. In late vision, unlike in pure beliefs, there can be no case of pure attribution, that is, of attribution of features in the absence of perceptually relevant particulars since the attributions are used to single out these particulars.

The concepts that figure in perceptual beliefs in late vision need not correspond to perceptual attributives, that is, they need not be restricted to concepts that late vision employs when it takes over the mode of presentation of the perceptual content. Visual systems have perceptual attributives for features such as shape, size, spatial relations, color, motion, orientation, texture, and affordances (Pylyshyn, 2003; Raftopoulos, 2009; Burge, 2010, p. 546), which are matched (partly, because one does not have concepts for all perceptual attributives) by the salient concepts. However, they do not have perceptual attributives for tigers, yet one does have perceptually based beliefs about tigers. They are perceptual in that

even though they do not conceptualize perceptual content and do not take over the mode of presentation of perceptions (category membership does not have a perceptual a mode of presentation), they depend for their empirical applications on perceptual attributives (the concept “tiger” depends for its application on perceptual attributives such as size, shape, and color).

I said that visual systems do not have perceptual attributives for category membership, which means that these higher-order properties cannot be visually represented; one does not perceive, say, tigerness, as Bayne (2009) and Siegel (2006) argue. Let me explain this. The fact that late vision outputs recognitional beliefs that are not pure beliefs does not entail that one has visual awareness of the high-level properties that figure in the recognitional beliefs. The IT cortex (which is the highest visual area) may represent objects in 3D, their 2D projections, viewer-centered representations, viewer independent representations, whole objects, and parts of objects, but not category membership. One has cognitive access awareness (CAA) of higher-level properties. (CAA is about perceptual content that is accessed by cognition becoming available to introspection and refers to episodes of thinking about the contents of one’s perceptual experience.) These beliefs are inextricably linked to a perceptual context but this does not entail that there is a visual phenomenology of category membership. It means, however, that the belief modulates top-down the processing in the visual areas of the brain and enhances the activation of the visible features that knowledge of the category membership highlights. Thus, having recognized an object affects the perception of some of its visible features by changing their representation and phenomenology, but one does not have visual awareness of high-level features of objects.

The inextricable link between thought and perception in late vision explains the essentially contextual, in Perry’s (2001) and Stalnaker’s (2008, pp. 78–82) sense, character of beliefs in late vision. The proposition expressed by the belief cannot be detached from the perceptual context in which it is believed, and cannot be reduced to some other belief in which some third person or objective content is substituted for the indexicals that figure in the thought (in the way one can substitute via Kaplan’s characters the indexical terms with their referents and get the “objective” truth-evaluable content of the belief). The reason is that the belief is tied to a idiosyncratic viewpoint by making use of the viewer’s physical presence and occupation of a certain location in space and time; the context in which an essentially indexical thought is believed is essential to the information conveyed. There are not, to paraphrase Stalnaker (2008, pp. 86–87), some relevant objective facts that the person (S1) who entertains the objective thought that purports to express the essentially indexical content has to learn in order to entertain the same content as S2 who uses the essentially contextual thought. This means that the way the world is thought by S2 is different from the way the world is thought by S1 not because there are some different facts the two thoughts are about, but because S1’s and S2’s perspectives on the same facts are different.

Perception individuates objects in a visual scene by assigning object-files based primarily on spatio-temporal information. The perception itself has the demonstrative reference force of “that object” and, thus, perceptual objects are determined relationally (Burge, 2010). For an object to be an object of a perceptual state

it must stand in a certain kind of relation to that state. Being acquainted perceptually with an object means that one is in direct contact with the object itself and retrieves information from it and not through a description (Burge, 1977). Perception puts one in a *de re* relationship with the object (as opposed to a *descriptivist* relationship). The content of an object-file is the *de re* mode of presentation of the object in perception (Raftopoulos and Muller, 2006).

Since recognitional beliefs rely on the presence of the object (reference to the object is fixed through a demonstrative as in “That x is F”), they are *de re* beliefs. Pure perceptual beliefs, on the other hand, have their referents fixed through a description of the object in memory. The *de re* relationship to a visual object eventually results in the formation of a *de re* belief about it. The outcome of late vision is a *de re* belief tied into a perceptual context. In contradistinction, pure thoughts and the pure attributions they render possible can be used outside any perceptual context and they descriptivist beliefs⁹.

It is sometimes argued that the main difference between thoughts and perceptions is that perceptual experiences, unlike thoughts, have a sensory quality to them (Dretske, 1993, p. 436). Although the amodal character of cognitive states as opposed to the modality-specific character of perceptions is a good place to start, this should be qualified because thoughts are not in a sense necessarily purely amodal since they may be accompanied by experiences that have a phenomenal character. The thought “the orange is round and yellow” has a modality-specific content, in that when one holds this thought, visual areas of one’s brain encoding color and shape are also be activated (Prinz, 2002). However, things are complicated. First, this activation does not entail that there is visual awareness of these features. Second there is a large literature on this issue with conflicting results. I am raising this issue to urge the reader not to take in views like Dretske’s uncritically.

BELIEFS: TAKE TWO

If the recognitional beliefs formed in late vision are not endorsed to become judgments, they are in some sense hypotheses. Suppose that upon viewing a scene containing an object O, S comes to believe that O is F. Since things may not be as they seem, S refrains from judging that O is F; S does not endorse the content of her perceptual belief. How is this recognitional belief different from the hypotheses or implicit beliefs that are constructed during the earlier stages of late vision in order to establish the identity of the object beyond the fact that the one is explicit, while the other is implicit?

In my view, the main difference consists in that the early hypotheses are tested against the iconic information stored in visual areas. This is an unconscious process that is outside the control of the viewer who is usually aware only of the content

of the winner, that is, the content of the explicit recognitional belief. However, the recognitional belief of late vision must be tested against a different sort of evidence in order to become a judgment. It must be tested against other sorts of propositional structures, that is, pure beliefs in which the predicate terms function as pure attributions. The aim of the testing is to put aside various possible defeaters of the belief. For example, the viewer has to decide whether she is the victim of some hallucination, etc. The processes involved in this testing may be available to the viewer’s consciousness, they are usually under her control, and they have the form of inferences from propositional contents to propositional contents, unlike the processes in late vision. The viewer tries to determine whether she should take the content of her late vision at face value. This is why testing the recognitional belief against other pure beliefs is a discursive process that is within the space of reasons, whereas testing the implicit hypotheses to come up with a recognitional belief belongs to late vision. In this sense, the recognitional beliefs formed in late vision are at the interface between the space of reasons and the perceptual space and, thus, have a pivotal role to play in accounts of justification of perceptual judgments.

I can explain now my claim that a belief is a dispositional state as opposed to a judgment that is an occurrent state. I tried to express the thought that perception gives us a *prima facie* inclination to believe that O is F but other evidence may override this and preclude us from forming the judgment that O is F. For example, some illusions give us a *prima facie* reason to believe that O is F but we do not endorse this because we do not believe that O is F. Undoubtedly, when O appears F in one’s experience, one is inclined to form (this is what I mean by “*prima facie*”) the recognitional belief that O is F. However, one need not endorse that thought. That O appears F in one’s experience should not be equated with one endorsing that O is F. To do that, one has to consider other relevant beliefs. Thus, to transform the belief to a judgment, one has to integrate it in the nexus of other beliefs, putting it, thus, within the space of reasons. This is possible because the recognitional belief already has a propositional structure.

There are two notions of belief here. The one is related to the expression of the content of a conceptual perceptual state, the recognitional belief, and the other is constitutively related to the notion of judgment. The relation of the belief in the first sense to late vision contents is not inferential. The relation of the same recognitional belief with the nexus of other beliefs is an inferential relation; if endorsed, the belief becomes a judgment. The belief is, thus, a disposition to make judgments (McDowell, 1994, p. 60), which do not introduce some new content but simply endorse the content of the recognitional belief.

Johnston (2006, p. 282) argues that the judgments that perception outputs are not inferentially based on perceptual content. “My judgment does not go beyond its truthmaker, which sensory experience has made manifest. Its truth is thus guaranteed by its origins. This is how immediate perceptual judgments often have the status of knowledge. There is no evidence from which they are inferred; instead they are reliable formed out of awareness of their truth maker, often in the absence of any evidence to the contrary.” Johnston talks about immediate perceptual judgments, whereas I talk about recognitional beliefs that may or may not

⁹In a *de re* belief, one retrieves information from the object itself and not through a description. In late vision where information in WM guides the formation of hypotheses about object identity, these hypotheses are based on descriptions in addition to visual information, since the knowledge stored in memory is a description of the object. Thus, the ensuing recognitional belief is based on a combination of information deriving from the object and from a description of it in memory. It is not a pure *de re* belief.

become judgments. Johnston's view that perceptual judgments are not inferred from perceptual evidence is correct. Our difference stems from considerations pertaining to the sentence "often in the absence of evidence to the contrary." I have claimed that to examine possible evidence against a recognitional belief, the belief must be inferentially tested against other pure beliefs (perceptual or otherwise). Only when it passes the test it becomes a judgment. Thus, I qualify Johnston's view that perceptual judgments are not inferred from any evidence, by distinguishing between perceptual beliefs and perceptual judgments and by adding that the former are not inferred from any evidence as outputs of late vision, but to become judgments they have to enter into inferential relations with possible defeaters.

I do not claim that recognitional beliefs are always tested this way to become judgments. Under normal conditions they are not tested at all. One might argue, however, that the absence of testing means that the viewer thinks that there is no reason to doubt the recognitional belief, which in itself is a sort of implicit inference. Or, one might think that in these normal cases, the recognitional belief becomes automatically a judgment without any inferential involvement. Still, the distinction holds because on certain occasions the recognitional belief is inferentially tested against other beliefs in order to become a judgment and, thus, recognitional beliefs and perceptual judgments belong to different categories, the first being a state that has the potential to become a judgment, even if the potentiality is actualized on certain occasions automatically.

LATE VISION AND AMODAL COMPLETION

Nanay (2010) thinks that mental imagery is necessary to account for amodal completion. He also (Nanay, 2010, p. 252) thinks that amodal completion in some cases is accompanied by some sort of phenomenology subserved by the activation of the early visual areas. In this sense, the hidden parts and features of an object are not merely believed in but are present in the object of perception as actualities by being imagined. Moreover, even in cases of amodal completion that are not accompanied by some sort of phenomenology, the hidden parts or features are perceptually represented. This is good point to delineate further the distinction between visual awareness and visual understanding and why late vision is a case of visual awareness. Briscoe (2011, pp. 165–167), argues that although imagery is sufficient for amodal completion, it is not necessary since one could either C-complete a visual scene by forming beliefs about the hidden parts of an object based on its visible features without projecting a mental image (the belief-based account of C-completion), or one could amodally complete a scene in bottom-up perceptual ways, in the way explained in the third section¹⁰.

Briscoe (2011) remarks that there are cases of C-completion, for example, the 3D sketch of an object whose backsides are hidden from view, which are cognitively driven in that to complete the hidden parts the viewer must draw from object knowledge. This may produce activation of the visual cortex, such that one has a mental image of the hidden parts, or it may produce simply

a thought that there are some parts hidden from view without any mental images, or it may produce both (Briscoe, 2011, p. 158). If the visual cortex is involved in C-completing the picture there is a synergy of bottom-up and top-down processes. 3D completion occurs in late vision where certain visual processing areas are activated.

If C-completion involves a pure perceptual thought about the hidden parts that results from an inference based on past experience and the current visual evidence, this is a case of visual understanding and not of visual awareness. I do not think that this possibility undermines my thesis that seeing the 3D sketch takes place in late vision. First, it is not clear whether there is empirical evidence for C-completion through pure thought and in the absence of any activation in visual areas. Second, if there are such cases, this only shows that sometimes C-completion does not occur in late vision but in discursive reasoning. Third, Briscoe's example from which he argues that C-completion may involve a pure thought involves a picture of the backside of what looks like a horse. In this case C-completion takes the form of a pure thought that this is a horse without any visual awareness. This is clearly a case of an inference involving visual understanding that occurs in the space of reasons and not in late vision. My claim is, on the other hand, that seeing the 3D sketch is a case of C-completion that takes place in late vision and involves visual awareness. Thus, even if there are cases of C-completions through pure thoughts, there are sorts of C-completions, such as seeing the 3D sketch, that take place in late vision and are cases of visual awareness.

Consider the white surface of a wall seen in a shadow and perceived as gray. Even though the viewer knows that the gray shade is caused by the shadow cast on a white wall, the phenomenal character of her experience is that of gray. The phenomenal character of her experience of the situation dependent color property (Schellenberg, 2008) or the phenomenal property (Shoemaker, 2006) is gray not white. Of course, being aware of the shadow she could infer the intrinsic (Schellenberg, 2008) or objective (Shoemaker, 2006) color of the wall but this is an inference based on the visible grayness, knowledge of the effects of shadows on surfaces, etc. In this case, one does not perceive the whiteness in any sense of "seeing" and, thus, the output of late vision is not the belief that the color of the wall is white. That the wall is experienced in late vision as gray is a case of visual awareness, where the concomitant belief takes over the mode of presentation of the object of experience. One may form the judgment that the wall is white even though it looks gray, but this representation is in the realm of pure thought. It is a case of visual understanding, a process in which one draws a conclusion based on the evidence of one's senses and other relevant information.

Suppose now that one sees one's hand moving back and forth. One sees the hand having the same size, a case of size constancy. If the constancy is due to cues that are available in the retinal image, the viewer is phenomenally aware of the same size despite differences in the viewing conditions. If size constancy is not effectuated through visual information and cognitive sources are needed, it is achieved in late vision; the viewer believes that the size is constant and has the phenomenal experience of a constant size. Should visual information be insufficient for perceptual constancies and should the non-visual information that ensures constancy be not

¹⁰Note that Nanay (2010, p. 244) seems to talk about a perceptually driven amodal completion that is insensitive to other beliefs.

available (as where attention is diverted elsewhere), the viewer would be aware of changes in size. This is what Epstein and Broota (1986) show by demonstrating that when attention is directed elsewhere, the size constancy operations fail. Thus, the experience of a stable size is the product of late vision, created by the knowledge of the size and stability of our hand in synergy with visual information coming from the hand. There is a large amount of literature supporting the view that many a perceptual constancy rely on object knowledge (Granrud, 2004; Cohen, 2008; Hatfield, 2009). Despite the role of thoughts in late vision, these cases should be better construed as visual awareness and not as visual understanding because, first, the states of late vision do not consist in pure thoughts but in hybrid states and, second, because the processes that lead to perceptual constancies are not discursive inferences.

To recapitulate, in pure thought the beliefs formed result from discursive processes (which may include perceptual information cast in a propositional form) and their attributives are context free, while in late vision there are no discursive processes but only conceptually modulated visual processing and the relevant attributives are context bound. These differences result from the constitutive involvement in late vision of visual circuits, an involvement that is absent in pure thought. This view entails that in amodal completion, which is one of the processes that take place in late vision, the missing or occluded features are not represented by pure perceptual beliefs, a view that is also supported by (partially) independent considerations offered by Nanay (2010, pp. 243–246).

REFERENCES

- Barr, M. (2009). The proactive brain: memory for predictions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1235–1243.
- Bayne, T. (2009). Perception and the reach of phenomenal content. *Philos. Q.* 39, 385–405.
- Biederman, I. (1987). Recognition by components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147.
- Briscoe, R. E. (2011). Mental imagery and the varieties of amodal perception. *Pac. Philos. Q.* 92, 153–173.
- Burge, T. (1977). Belief de re. *J. Philos.* 74, 338–362.
- Burge, T. (2010). *Origins of Objectivity*. Oxford: Clarendon Press.
- Chelazzi, L., Miller, E., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345–347.
- Cohen, J. (2008). Colour constancy as counterfactual. *Australas. J. Philos.* 86, 61–92.
- Delmore, A., Rousselet, G. A., Mace, M. J.-M., and Fabre-Thorpe, M. (2004). Interaction of top-down and bottom up processing in the fast visual analysis of natural scenes. *Brain Res. Cogn. Brain Res.* 19, 103–113.
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
- Dretske, F. (1993). Conscious experience. *Mind* 102, 263–283. [Reprinted in Noë and Thompson (2002). *Vision and Mind*. Cambridge, MA: MIT Press].
- Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Epstein, W., and Broota, K. D. (1986). Automatic and attentional components in perception of size-at-a-distance. *Percept. Psychophys.* 40, 256–262.
- Fabre-Thorpe, M., Delorme, A., Marlot, C., and Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *J. Cogn. Neurosci.* 13, 171–180.
- Granrud, C. E. (2004). “Visual metacognition and the development of size constancy,” in *Thinking and Seeing*, ed. D. T. Levin (Cambridge, MA: MIT Press), 75–95.
- Hatfield, G. (2009). *Perception and Cognition*. Oxford: Clarendon Press.
- Jackendoff, R. (1989). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.
- Jackson, F. (1977). *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- Johnson, J. S., and Olshausen, B. A. (2005). The earliest EEG signatures of object recognition in a cue-target task are postsensory. *J. Vis.* 5, 299–312.
- Johnston, M. (2006). “Better than mere knowledge: the function of sensory awareness,” in *Perceptual Experience*, eds T. S. Gendler and J. Hawthorne (Oxford: Clarendon Press), 260–291.
- Kihara, K., and Takeda, Y. (2010). Time course of the integration of spatial frequency-based information in natural scenes. *Vision Res.* 50, 2158–2162.
- Kosslyn, S. M. (1994). *Image and Brain*. Cambridge, MA: MIT Press.
- Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends Cogn. Sci. (Regul. Ed.)* 7, 12–18.
- McDowell, J. (1994). *Mind and World*. Cambridge, MA: Harvard University Press.
- McPherson, F. (in press). Cognitive penetration of colour experience: rethinking the issue in light of an indirect mechanism. *Philos. Phenomenol. Res.*
- Nanay, B. (2010). Perception and imagination: amodal perception as mental imagery. *Philos. Stud.* 150, 239–254.
- Palmer, S. (1999). *Vision Science*. Cambridge, MA: MIT Press.
- Perry, J. (2001). *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
- Peyrin, C., Michel, C. M., Schwartz, S., Thut, G., Seghier, M., Landis, T., Marendaz, C., and Vuilleumier, P. (2010). The neural processes and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. *J. Cogn. Neurosci.* 22, 2678–2780.
- Prinz, J. J. (2002). *Furnishing the Mind*. Cambridge, MA: MIT Press.
- Pylyshyn, Z. (2003). *Seeing and Visualizing: It's not what you Think*. Cambridge, MA: MIT press.
- Raftopoulos, A. (2009). *Cognition and Perception: How do Psychology and the Neural Science Inform Philosophy?* Cambridge, MA: MIT Press.
- Raftopoulos, A. (2010). Can non-conceptual content be stored in visual memory? *Philos. Psychol.* 23, 639–668.
- Raftopoulos, A., and Muller, V. (2006). Nonconceptual demonstrative reference. *Philos. Phenomenol. Res.* 72, 251–285.

- Rock, I. (1983). *The Logic of Perception*. Cambridge, MA: MIT Press.
- Roelfsema, P. R., Lamme, V. A. F., and Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395, 376–381.
- Schellenberg, S. (2008). The situation dependency of perception. *J. Philos.* 105, 55–84.
- Shams, L., and Beierholm, U. R. (2010). Causal inference in perception. *Trends Cogn. Sci. (Regul. Ed.)* 14, 425–432.
- Shoemaker, S. (2006). “On the way things appear,” in *Perceptual Experience*, eds T. S. Gendler and J. Hawthorne (Oxford: Clarendon Press), 461–481.
- Siegel, S. (2006). “Which properties are represented in perception?” in *Perceptual Experience*, eds T. S. Gendler and J. Hawthorne (Oxford: Clarendon Press), 481–504.
- Spelke, E. S. (1988). “Object perception,” in *Readings in Philosophy and Cognitive Science*, ed. A. I. Goldman (Cambridge, MA: MIT Press), 447–461.
- Stalnaker, R. C. (2008). *Our Knowledge of the Internal World*. Oxford: Clarendon Press.
- Strawson, P. (1974). “Imagination and perception,” in *Freedom and Resentment*, ed. P. Strawson (London: Methuen), 45–65.
- Treisman, A. (2006). How the deployment of attention determines what we see. *Vis. cogn.* 14, 411–443.
- Treisman, A., and Kanwisher, N. G. (1998). Perceiving visually presented objects: recognition, awareness, and modularity. *Curr. Opin. Neurobiol.* 8, 218–226.
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 12 August 2011; accepted: 01 December 2011; published online: 16 December 2011.
- Citation: Raftopoulos A (2011) Late vision: processes and epistemic status. *Front. Psychology* 2:382. doi: 10.3389/fpsyg.2011.00382
- This article was submitted to *Frontiers in Perception Science*, a specialty of *Frontiers in Psychology*.
- Copyright © 2011 Raftopoulos. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.