

Timing Time: Why Early Vision is Cognitively Impenetrable

Athanasios Raftopoulos (raftop@ucy.ac.cy)
Department of Psychology, University of Cyprus, P.O. Box 20537
Nicosia 1678, Cyprus

Abstract

Newen and Vetter (2016) argue that cognitive penetration (CP) of perceptual experience is the most possible account of the evidence. They target both the weak impenetrability thesis that only some early visual processes are cognitively impenetrable (CI), and the strong impenetrability thesis that all perceptual processes are CI. Since I agree that perceptual processing as a whole is CP, I will concentrate on their arguments against the weak CI thesis. In attacking weak CI, the authors take aim at Raftopoulos' arguments supporting the CI of early vision. Their main argument comes from studies that, Newen and Vetter think, show that early vision is CP by demonstrating the existence of cognitive effects on early vision. I examine the same empirical evidence that Newen and Vatter discuss and argue this same evidence strongly supports the view that early vision is CI.

1. Introduction

Newen and Vetter (2016) argue that the CP of perceptual experience is the most possible account of the available evidence. They target both the weak impenetrability thesis that only some early visual processes are CI, and the strong impenetrability thesis that all perceptual processes are CI. Since I agree that perceptual processing, as a whole, is CP, I concentrate on weak CI. The authors criticize Raftopoulos' (2001, 2014) arguments in support of the view that early vision, defined by temporal criteria, is CI. They maintain that top-down influences from higher-level cognitive areas to early visual cortex occur very early. Thus, it is unjustified to conclude from an early timing of a visual process that it is unaffected by top-down cognitive influences.

The authors' main argument against weak CI, given that weak CI holds that during the timing of early vision no direct cognitive effects modulate the ongoing perceptual processes, comes from studies that, Newen and Vetter think, demonstrate the existence of cognitive effects on visual processing that occur within the time frame of early vision.

Furthermore, Newen & Vetter (2016, 5) argue that even if it exists a stage of perceptual processing that is unaffected by cognition, it lasts for so few ms that it could not be a plausible candidate for a perceptual module and it would be of almost no importance to philosophical discussions about the CP of perception. In other words, even if such a CI stage exists, the information it processes and outputs would be too poor to be categorized as a properly speaking perceptual content.

Here, I concentrate on Newen and Vetter's arguments from temporal processing. The main reason is that if it turns out that this evidence does not support the CP of early vision, the rest of the arguments against weak CI are moot.

I argue that the evidence Newen and Vetter adduce if properly examined strongly supports weak CI. The evidence that emerges from these studies shows that all visual brain areas at some time are affected by top-down cognitive signals that usually drive spatial or object/feature based attention. They also show that within the time frame of early vision there is a confluence of top-down, lateral, and bottom up interactions. These interactions, however, do not involve any cognitive signals. These studies do not show that during the first 120-140 ms of perceptual processing there are direct cognitive top-down effects on early vision. What Newen and Vetter consider as evidence for top-down cognitive effects on early vision is, in effect, evidence for top-down and bottom up interactions that do not involve any cognitive effects.

In the first section, I define early vision and examine most of the studies cited by Newen and Vetter (2016) and argue that they do not show that early vision is CP. I also take up Newen and Vetter's claim that even if a CI stage of visual processing exists it is too impoverished to be deemed a perceptual stage and claim that the output of early vision has a rich structure. In the second section, I examine the rest of the evidence used by Newen and Vatter and argue that in effect it supports the claim that early vision is CI. All recurrent processing during early vision is restricted within the visual areas of the brain and does not involve any causal influences from cognitive states.

2. Early Vision and why it is CI, Part 1.

Early vision includes a feed forward sweep (FFS) in which signals are transmitted bottom-up. In visual areas (from LGN to FEF) FFS lasts for about 100ms. Early vision 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 80–100 ms and culminates at 120–150 ms. Lamme (2003) calls it local recurrent processing (LRP). The unconscious FFS extracts high-level information that leads to categorization, and results in some initial feature detection. LRP produces further binding and segregation. The LRP is needed because,

owing to the small receptive fields of the neurons in V1 and V2, only local information can be coded at this level. The segmentation and recognition of the objects in a visual scene requires a more global analysis of the visual scene that can be achieved in higher areas, such as V4 or MT/V5, where the neurons' receptive fields are larger and integrate information across longer distances in the visual field.

The feedback projections provide this global analysis that allows object segmentation, figure/ground separation, and object recognition. In the case of the MT/V5 feedback to V1, there is evidence (Pomp et al, 2015) that this feedback increases the responsiveness of the neurons in V1 especially for low-salience, small signals, which means that the recurrent signals from MT/V5 may serve to disambiguate sub-optimal visual input with respect both to the spatial location and motion of the sub-optimal signals and to their content. In addition, the feedback signals may be used to inform V1 where a change has happened in the visual scene. By not involving signals from the cognitive areas of the brain, FFS and LRP are cognitively impenetrable since the transmission of signals within the visual system is not affected by top-down signals produced in cognitive areas.

The processes of early vision retrieve from the environment information that allows the perception of a visual scene with as much accuracy as possible. To do so, early vision gradually constructs representations of increasing complexity (from variations in light intensities it extracts edges, from edges blobs, from blobs it extracts two-dimensional surfaces, and from these the 2 $\frac{1}{2}$ sketch). The output of early vision consists in the dynamic structural description of a visual scene corresponding to Marr's 2 $\frac{1}{2}$ sketch to which one could add the affordances of objects.

Current research on predictive coding sheds light on the nature of the processes implicated in vision. Applying this to early vision, one gets the following. The top-down and lateral effects in early vision aim to test hypotheses concerning the putative distal causes of the sensory data encoded in the hierarchically lower neuronal assemblies. In this testing, predictions made on the basis of hypotheses about the sensory information that the lower levels should encode assuming that the hypotheses are correct, are tested against the actual sensory information encoded at the lower levels. The hypothesis that best matches the sensory data is selected.

To form hypotheses concerning the probable cause of the sensory data at a certain level, at a specific spatial, and temporal scale, the neuronal assembly at the next level uses information not only about the sensory data at the previous level (or, to be precise, information regarding its prediction error) that is transmitted bottom-up, but also higher-level information that is transmitted either laterally, that is, from neuronal assemblies at the same level (neurons in V1 processing wave-lengths inform other neurons in V1 processing shape information), or top-down from levels higher

in the hierarchy (neurons in V4 are informed about the color of incoming information from neurons in IT). This lateral and top-down processing provides the context in which each neuronal assembly constructs the most probable hypothesis that would explain the sensory data at the lower level.

Since 90% of the information transmitted by neurons is transmitted within the first 100 ms of the neurons' activation as a response to a stimulus, information to neurons transmitted from other assemblies can affect their activity only if it arrives within the 100 ms time frame (Bullier 2001, 98). In order for the recurrent signals to modulate the activity at the reentered sites, they should reenter them during these crucial 100ms. Thus, for some signal from V4 or MT/V5, which receive feedforward signals from V1, to reenter V1 in time to influence the activation of the V1's neurons, the loop consisting of feedforward signals from V1 to V4 or MT/V5 and the recurrent signals from V4 or MT/V5 back to V1 must have been completed in less than 100ms.

To put things into perspective, let us revisit Bullier's (2001) 'reintroduction' view as it pertains to early vision and involves MT/V5 and its interaction with the lower visual areas V1 and V2. Low spatial frequency (LSF) signals precede high spatial frequency (HSF) signals. LSF information is transmitted through fast magnocellular pathways, while HSF information is transmitted through slower parvocellular pathways; the information transmitted through M-channels reaches V1 from LGN 20 ms earlier than the information transmitted from LGN to V1 through P-channels.

The mean activation latency of the neurons in MT/V5 of the brain is 75 ms after stimulus presentation respectively. Signals arrive at these areas at about the same time as, or a bit later than, they arrive in V1 (50-80 ms) and V2 (85 ms) and much earlier than they arrive in V4 despite the fact that MT/V5 is anatomically higher than V4 (Bullier (2001, 98). MT/V5 (and FEF) are parts of the 'fast brain' and belong to the dorsal system. MT/V5 is situated in the parietal cortex. Signals from V1 can reach the MT/V5 at about the same time they reach V2, that is, within 1-2 ms. It takes less than 20 ms for the recurrent signals from MT/V5 to affect the activation of neurons in V1 and V2. So, when HSF information transmitted through the P-channels reaches V1, 20 ms after LSF information transmitted through M-channels had reached V1, the responses of the V1 neurons have been modified as a result of the top-down signals from MT/V5 that had received earlier LSF information. In addition to the fast transmission of signals through the M-channels, MT/V5 also receives fast signals directly from LGN bypassing V1 through the koniocellular pathway. Thus, under certain conditions, MT/V5 could be activated earlier than V1 (Fuxe & Simpson 2002).

Bullier (2001, 100) concludes "the first wave of activity that invades the visual cortex following a visual stimulus appearing in the visual scene is carried by the M channel . . . the characteristics of the M channel are well suited for such a

first-pass analysis of the visual scene.” This entails that even the earliest ERP component, C1, which is elicited between 40-60 ms, is not an indice of the activity of V1 alone but also likely reflects top-down influences to V1 from areas as high as MT/V5. These bottom-up and top-down interactions take place within early vision (they have latencies up to 140 ms), but all these studies bear evidence to the existence of top-down flow of information within early vision that involves stimulus driven signals reprocessed in higher visual areas and no evidence for cognitive effects in early vision.

The reason that the picture described thus far bears no evidence to support the existence of cognitive influences at these early latencies is that the top-down signals are transmitted from MT/V5 and are part of the processing along the dorsal system. The picture presented thus far posits early latencies of the signals arriving to MT/V5 from LGN either directly or through V1, that is, it dictates a bottom-up early activation of MT/V5 that, in turn, sends feedback to V1 affecting the activations of the neurons there. At these latencies there are no top-down cognitive signals to MT/V5 and, therefore, there are no cognitive signals affecting V1.

This conclusion is reinforced by the results of a study by Plomp et al (2015). The researchers combined EEG source-imaging and Granger-causal modeling with high temporal resolution to “investigate whether recurrent and top-down interactions between visual and attentional brain areas can be identified and distinguished at short latencies in humans.” Their results confirm the fast interaction between V1 and MT/V5 reported by Bullier (2001). Their results also show that the C1 ERP waveform (50 ms (onset)-80/90 ms (peak)) reflects both V1 activity and also activity in highly distributed areas situated at the occipital, parietal, and frontal lobes (FEF is in the pre-frontal lobe in an area where the dorsal pathway projects). This activity reflects the early bottom-up and top-down interactions described by Bullier that includes the V1/MT feedback loop. In contradistinction to this early recurrent activity, the parietal cortex and FEF (the later cycle of activity there), which are known to modulate perceptual processing so as to help adapt behavior to the demands of a task and context, affect posterior activity around the latency of N1 (170 ms after stimulus onset). Thus, top-down interactions that reflect task-specific processing of the stimuli arise at longer latencies after stimulus onset. Pomp et al. (2015, 4-5) synopsise their results as follows “at the N1 latency, driving from MT no longer showed a stimulus effect, indicating that stimulus-specific driving from MT is confined to earlier latencies, in line with its fast response properties.”

As Plomp et al., (2016, 1) write “stimulus-evoked activity at latencies before 100 ms is traditionally considered a bottom-up process. Even at these short latencies, however, there is mounting evidence of fast recurrent interactions between visual areas, obtained from direct recordings of neural activity in animal models.” This agrees with Bulier’s

conclusion that during the early interaction between MT/V5 and V1 (latencies earlier than 100 ms) the signals are stimulus driven (since the signals entering MT and processed there originate from the stimulus only), while only in the later interactions that involve cognitively driven attention, whose commands are issued according to the task demands, do cognitive factors modulate the activation of neurons in V1.

Suppose that early vision is CI. Newen and Vetter could be right that the perceptual processes within this narrow time frame produce states with such poor contents that they are not properly speaking perceptual states; they could be at most sensory states. As we have seen, however, FFS and local RP allow, in about 120-140 ms after stimulus onset, the construction of fairly complex representations of stimuli. There is some form of perceptual organization, which certainly includes information regarding the presence of discrete objects in a scene, their orientations, sizes, shapes or forms, motions; these features determine the structural description of objects. Thus, the output of early vision consists of information about spatio-temporal and surface properties, 3D shape viewed from the perspective of the viewer, 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 (Raftopoulos 2014). I disagree, thus, with Newen and Vetter that the content of CI perceptual states is related to the perception of impoverished black and white pictures. Early vision retrieves from the visual scene an extensive range of information.

3. Timing the Cognitive Effects: Why Early Vision is CI, Part 2

Let us examine the evidence that Newen & Vetter employ to substantiate their claim.

Time-resolving electrophysiological evidence showed that visual cortex is activated within 50 ms and pre-frontal areas within 80 ms after visual stimulus onset. This leaves plenty of time for iterative top-down processing between “cognitive”, e.g. frontal and parietal, areas and sensory, e.g. occipital, areas, within the first 100–200 ms after visual stimulation (Fuxe & Simpson, 2002). Thus, complex high level and reiterative processing can happen very fast and can influence visual processing very early on (. . . Plomp, Hervais-Adelman, Astolfi, & Michel, 2015). (Newen & Vetter 2016, 4-5)

Newen and Vetter talk about recurrent signals that involve cognitive activity affecting visual areas at latencies 100-200 ms. Thus, they accept that the available evidence suggests that cognitive effects on visual areas are registered after 100 ms post stimulus. We discussed Plomp et al. (2015) work and their conclusion that the recurrent early activity (before the elicitation of N1) is restricted within visual areas and only after that latency does recurrent activity involving cognitive

centers register in visual areas. Thus, the early recurrent activity (up to about 170 ms) does not involve cognitive signals and this supports the view that the processes of early vision (that lasts up to about 120-140 ms) are not affected by cognitive signals. Thus, when Newen and Vetter conclude “complex high level and reiterative processing can happen very fast and can influence visual processing very early on”, this very early on is not early enough to be within early vision.

Indeed, as Foxe and Simpson (2002, 139) state

There is clearly sufficient time for multiple iterations of interactive processing between sensory, parietal, and frontal areas during brief (e.g., 200 ms) periods of information processing preceding motor output . . . These data strongly suggest that activity represented in the “early” ERP components such as P1 and N1 (and possibly even C1) is likely to reflect relatively late processing, after the initial volley of sensory afference through the visual system and involving topdown influences from parietal and frontal regions.

Notice, first, that the reference is to the time frame up to 200ms and that the recurrent interactions at earlier latencies that Foxe & Simpson report concern interactions within visual areas. In addition, the top-down signals that are generated in the higher visual areas and reenter the early visual areas within these earlier latencies result from the processing of sensory signals that arrive very quickly, through PM-channels or the koniocellular pathway, to the higher areas; “The rapid activation of prefrontal cortex following initial visual activation (within 30 ms) suggests that this input is mediated through the faster dorsal visual stream” (Foxe & Simpson 2002,147-148). This is in line with Bullier’s views that we examined earlier. All these suggest that the higher areas at these latencies have not received as yet any signals from cognitive areas and, in this sense the signals that constitute the feedback loop are bottom-up sensory signals and top-down reprocessed and modified sensory signals. In fact, it could hardly be otherwise; since all this very early recurrent activity involves the dorsal system, there is up to date no evidence to support the existence of any cognitive effects on the dorsal system when it functions on line to support fast action.

Furthermore, and in reference to Foxe and Simpson’s mention of parietal and frontal regions involved in the early recurrent processing, which may be taken as evidence for the existence of cognitive influences, MT/V5 is in the parietal cortex and FEF is in the prefrontal cortex. Our discussion concerning the role of MT/V5 shows that there are no cognitive effects in the early latencies we discuss and, as our examination of the role of FEF will show, neither are such cognitive effects found in the early activation of FEF and in its role in the early stages of perceptual processing. Foxe and Simpson (2002, 146) confirm this analysis by concluding that “multiple *visual areas* begin to contribute substantially to the surface potential and C1 begins to reflect contributions from a

number of *visual areas* other than, but is likely also to include V1 (emphasis added).” Moreover, Foxe and Simpson (2002, 147), after their claim that “that sustained activation patterns within cortical areas are consistent with feedback modulation of ‘lower’ visual areas by ‘higher’ areas, as well as local intrinsic processing”, add that their findings conform with the findings of Lamme (1995) and Lamme et. al. (1998) about the time frame of feedback modulation in figure-ground segregation studies with monkeys. It is well known that these studies confirm that the recurrent processes that occur at early latencies do not involve cognitive signals.

Newen and Vetter (2016, 5) argue that in the visual system there is strong evidence for fast top-down processing within the first 50 ms after stimulus onset, certainly between motion area V5 and primary visual cortex V1 during motion perception (Silvanto, et al., 2005). The reference to the interaction between V5 and V1 during motor perception brings into mind the foregoing discussion of LRP that is restricted within the visual areas. Let us examine these studies to see if this assumption is substantiated.

Silvanto et al., (2005) studied the role of V1 in the visual awareness of motion. Their experiments show that back-projections from extrastriate cortex influence the activations of neurons in V1 and that it is the activation in V1 that determines which information reaches awareness. Since our interest is in the latencies at which the back projections affect V1 and the sites of origin of the top-down signals, I will ignore the findings concerning motion awareness. Silvanto et al., (2005) applied TMS on V1 and V5 at different times to examine the perception of phosphenes. When subthreshold TMS (that is, TMS producing no phosphene on its own) was applied over V5 followed by a subthreshold pulse to V1, subjects did not report any phosphene. When a subthreshold pulse was applied over V5 followed 10–40 ms later by a suprathreshold pulse over V1, subjects reported a phosphene, which was not merely the suprathreshold V1 phosphene. Instead, it acquired features of a suprathreshold V5 phosphene since subjects reported the perception of movement, and the shape and size of their percept was a mixture of V1 and V5 phosphenes. This shows that activity in V5, which on its own is insufficient to induce a moving percept, can produce such a percept if the level of induced activity in V1 is high enough.

Silvanto et al. (2005, 143) conclude that the fact “that moving phosphenes are perceived only when suprathreshold V1 stimulation follows, but not precedes, subthreshold V5 stimulation, together with the gradual increase in motion perception from the 10–50 ms period, precludes a simple feed-forward summation account and points instead to a critical time of backprojection arrival in V1.” They also note that the narrow time window for V5–V1 interaction (10–50 ms) is consistent with previous reports of extrastriate-striate feedback interactions in motion during this time interval. Indeed, this accords with Bullier’s (2001) and Plomp et al. (2016) finding

that there is an early (up to 100 ms) phase of recurrent activity between V1 and MT/V5, but as in these studies, so in Silvanto et al. (2005) report, there is no evidence to suggest top-down cognitive effects at these early latencies, because the recurrent signals from MT/V5 are stimulus driven, or, to use Plomp et al. (2016) term, they are a stimulus-evoked activity.

Next, Newen and Vetter (2016, 5) examine the interaction between FEF and V1.

[T]he frontal eye fields (FEF), a higher-level area in frontal cortex involved in motor planning of eye movements, exerts its influence to V5 within 30 ms (Silvanto, Lavie, & Walsh, 2006). Therefore, a feedback loop from a frontal region to an early occipital region can take as little as 80 ms or less . . . when the task requires face recognition, FEF signals are sent to face-sensitive regions and when the task requires motion discrimination, FEF signals are sent to motion area V5, both within a time frame of 20–40 ms after FEF activity (Morishima et al., 2009).

FEF is situated in the prefrontal cortex at a site that is heavily interconnected with the parietal cortex and is considered a part of the dorsal system. The mean activation latency of the neurons in FEF is 70 ms after stimulus presentation. Signals arrive at FEF with a slight time delay time with respect to the signals arriving at V1 (50-80 ms) and V2 (85 ms) and much earlier than they arrive at V4 despite the fact that FEF is anatomically higher than V4 (Taylor and Nobre 2007). FEF TMS affects the detection of targets in arrays of distractors and these effects are apparent when pulses are applied early (40 and 80 ms) after presentation of the visual array (O’Shea et al. 2004). HSF signals from V1 can reach FEF in 50-100 ms.

FEF contains visual and movement neurons. Studies (see O’Shea et al. (2004, 1060) for a discussion) show that there are two dissociated processing operations in FEF; the target selection by FEF visual neurons and saccade programming by movement neurons. Studies (O’Shea et al., 2016; Silvanto et al., 2006; Taylor & Nobre 2007) show that some FEF responses are independent of saccades to targets and respond to the visual stimuli. Some of the FEF feedback signals play a role in the perception of a visual scene by affecting in a top-down manner the earlier visual areas. FEF plays a crucial role in visual target discrimination that is independent of saccade programming, as TMS applied to FEF impairs performance in target discrimination tasks if applied between 40-80 ms after stimulus onset (O’Shea et al. 2004). In addition, these visual neurons of FEF are thought to be associated with top-down or endogenous attention (Taylor and Nobre 2007).

Accepting O’Shea et al. (2004) early latencies of FEF neurons in discriminating targets from non targets (100-120 ms), in view of the fact that, as Newen and Vetter (2016, 5) also accept, FEF exerts its influence to V5 within 30 ms and, therefore, a feedback loop from FEF to an early occipital

region can take as little as 80 ms or less, the total time it takes for the FEF neurons that have distinguished the targets from non targets to affect via top-down feedback projections the early visual areas is about 180-200 ms, considering that the target discrimination in FEF reported by O’Shea et al. (2004) occurs at 100-120 ms. This means that the FEF effects the activation of the neurons in early visual areas with a latency that places these effects outside early vision.

Concerning the finding that FEF neurons effectively discriminate targets from non targets as early as 100-120 ms after stimulus onset, one could argue that since this discrimination is task relevant and involves cognitive factors, cognition affects a visual area, FEF, within the timing of early vision. O’Shea et al. (2004), think it very likely that the early latency they report is the result of feature pre cueing, which means that the early activity in FEF occurs as the result of a cognitive demand issued before the appearance of the stimulus. I have argued (Raftopoulos 2014) that the cognitive effects on perception through pre-cueing are not cases of CP because they do not affect directly early vision and do not affect its epistemic role in grounding empirical beliefs.

Silvanto et al., (2006), whose study is cited by Newen and Vetter (2016) as showing that early vision is CP, found that stimulation applied to FEF 20-40 ms prior to the stimulation of MT/V5 decreases the intensity of the MT/V5 stimulation required to elicit phosphenes, which entails that the activity of MT/V5 is modulated by the activity in FEF. FEF has also been found to modulate top-down V4. Silvanto et al (2006, 944) claim that the content of to-down control may be either spatial or feature related, which means that they think that FEF affects the control of top-down attention, “an area involved in control would be expected to be active early and by responding to target features, the FEF could increase the sensitivity of extrastriate neurons to task relevant parameters.” (Silvanto et al. 206, 944) With regards to how FEF exerts top-down control, it is possible that FEF activity occurs prior to sensory stimulation as opposed to rapid responses to visual stimuli since FEF neurons may also play a role in visual priming (Silvanto et al. 2006, 944). Thus, as Taylor and Nobre (2007), so Silvanto et al., (2006) think that FEF controls the allocation of top-down attention prior to stimulus presentation.

The discrimination between targets and non-targets depends on the task at hand and is cognitively driven. Thus, the top-down effects that result from this discrimination are also cognitively-driven and the visual processes that are thus affected are clearly CP. Accepting O’Shea et al. (2004) early latencies of FEF neurons in discriminating targets from non targets, since FEF exerts its influence to V5 within 30 ms and, therefore, a feedback loop from FEF to an early occipital region can take as little as 80 ms, the total time it takes for the FEF neurons that have distinguished the targets from non targets to affect via top-down feedback projections MT/V5 is 130-150 ms, and the effects on the early visual areas is about

180-200 ms, considering that the target discrimination in FEF in O'Shea et al. (2004) occurs at 100-120 ms. All of these are outside the timing of early vision and do not entail its CP.

Finally, Newen and Vetter (2016, 5), appeal to a study by Drewes et al. (2016) that shows that object recognition involves recurrent processing with a time constant of 60 ms. Drewes et al. (2016) examine the view that since the visual system extracts from object information, for example, the shape of objects, very fast this entails that the underlying cortical processing should be strictly feedforward. Against this, their study suggests that in shape perception there is a recurrent circuit, which is not an attentional cueing effect but reflects "the time course of feedback processing underlying the rapid organization of shape." (Drewes et al. 2016, 185)

In their introduction, they mention work by Heinen et al. (2005) suggesting that the figure-ground segregation requires two distinct periods of information processing in the early visual areas, an early one around 130-160 ms and a later one around 250-280 ms after stimulus onset, and by Wokke et al. (2012) showings that recurrent processing engages the early visual areas (V1/V2) to participate in more complex visual tasks. In an early time window (96–119 msec), detection of figure stimuli and of neural correlates of figure border detection and border ownership occurs. Later (236–259 msec) V1 and V2 participate in surface segregation. Drewes et al. (2016) accept these latencies as a general framework.

Drewes et al. (2016, 190) claim that "the extent of facilitation between two shape stimuli depends non-monotonically on the delay between their presentations, peaking at a delay of 60 ms." This suggests a recurrent circuit underlying shape processing in the cortical object pathway. They remark that in Wokke's et al. (2012) study TMS was applied to the occipital pole to disrupt processing in V1/V2 or to the lateral occipital lobe to disrupt processing in the LOC. TMS disrupted performance at both locations but at different latencies. In LOC, TMS disrupted processing when the pulse occurred 100–122 ms post stimulus, while in V1/V2, processing was disrupted when the pulse was applied 160–182 ms post stimulus. This shows a feedback process in the grouping of contour fragments to form shape with a one-way feedback time constant (LOC to V1/V2) of 40–80 ms. Given the 60 ms time constant, the top-down signals reenter V1 and V2 at latencies outside early vision.

4. Conclusion

I examined the evidence Newen and Vetter (2016) adduce to support the claim that early vision is CP. None of it supports the existence of direct cognitive effects on early vision. Finally, concerning the claim that a stage of visual processing that is CI is so impoverished that it would not be worthy to be called a stage of perception, I claimed that early vision delivers a rich structure.

References

- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, 36, 96-107.
- Drewes, J., Goren, G., Zhu, W., & Elder, J/H. (2016). Recurrent processing in the formation of percept shapes. *The Journal of Neuroscience*, 36(1), 185-192.
- Heinen, K., Jolij, J., & Lamme, V. (2005). Figure-ground segregation requires two distinct periods of activity in V1: a transcranial magnetic study. *Neuroreport*, 16(13), 1483-1487.
- Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, 15, 1605-1615.
- Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 7 (1), 12-18.
- Lamme VA, Zipser K, Spekreijse H (1998). Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings National Academy Science USA*, 95, 3263–3268.
- Morishima, Y., Akaishi, R., Yamada, Y., Okuda, J., Toma, K., & Sakai, K. (2008). Task-specific signal transmission from prefrontal cortex in visual selective attention. *Nature Neuroscience*, 12(1), 85-90.
- Newen, A. & Vetter, P. (2016). Why cognitive penetration of our perceptual experience is still the most plausible account. *Consciousness and Cognition*, <http://dx.doi.org/10.1016/j.concog.2016.09.005>.
- O'Shea, J., Muggleton, N.G., Cowey, A., & Walsh, V. (2004). Timing of target discrimination in human front eye fields. *Journal of Cognitive Neuroscience*, 16(6), 1060-1067.
- Plomp, G., Hervais-Adelma, A. Astofli, L., & Michel, C. M. (2015). Early recurrence and ongoing parietal driving during elementary visual processing. *Nature Scientific Reports*, 5:18733, doi: 10.1038/srep18733.
- Raftopoulos, A. (2001). Is perception informationally encapsulated? The Issue of the Theory-Ladeness of Perception. *Cognitive Science*, 25, 423-451.
- Raftopoulos, A. (2014). The cognitive impenetrability of the content of early vision is a necessary and sufficient condition for purely nonconceptual content. *Philosophical Psychology*, 27 (5), 601-620.
- Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (V1) activity gates awareness of motion. *Nature Neuroscience*, 8(2), 143-144.
- Silvanto, J., Lavie, N., & Walsh, V. (2006). Stimulation of the human frontal eye fields modulates sensitivity of extrastriate visual cortex. *Journal of Neurophysiology*, 96(2), 941–945.
- Taylor, P.C.J., & Nobre, A. (2007). FEF TMS affects visual cortical activity. *Cerebral Cortex*, 17, 391-399.
- Wokke, M., Sligte, I., Scholte, H., & Lamme, V.A.F. (2012). Two critical periods in early visual cortex during figure-ground segregation. *Brain and Behavior*, 2(6), 763-777.