The Cognitive Penetrability of Perception

New Philosophical Perspectives

EDITED BY
John Zeimbekis
and
Athanassios Raftopoulos

OXFORD UNIVERSITY PRESS
The Cognitive Penetrability of Perception
An Overview

Athanassios Raftopoulos and John Zeimbekis

The hypothesis that perception is cognitively penetrable holds that cognitive states such as beliefs, desires, and possibly other states can causally influence perceptual processing in such a way that they end up determining subjects' perceptual contents or experiences. The philosophical significance of penetrability is easy to grasp: if perception is cognitively penetrable, then what we think literally affects how we see the world. This book elucidates the nature of the cognitive penetrability hypothesis, assesses its plausibility, and explores what the philosophical consequences would be if perception turned out to be cognitively penetrable or cognitively impenetrable.

Section 1 provides a brief outline of the origins and history of the cognitive penetrability debate. Section 2 disambiguates key concepts and explains their roles in the debate: impenetrability, modularity, and informational encapsulation; early vision and late vision; hard-wired perceptual formation principles; perceptual learning; theory-ladenness of observation; the role of attention in cognitive penetration; and the perception/cognition distinction itself. Section 3 presents existing definitions of the cognitive penetrability of perception. Section 4 explains the relevance of the penetrability debate to a range of philosophical topics: the contents of perception; cognitive phenomenology and epistemic feelings; nonconceptual content; awareness and attention; perceptual warrant; realism and representationalism; and action. The contributed chapters are summarized in the sections of the introduction that deal with the corresponding topics.

1 We cross-refer to sections within the Introduction by e.g. 'Section 3' or simply 'see 3'. Where we refer to a section in another chapter in the volume, we use the symbol §.

2 We often abbreviate 'cognitive penetrability of perception' and 'cognitive impenetrability of perception' to 'penetrability' and 'impenetrability', and sometimes to 'CP' and 'CI.'
1 Cognitive Penetrability: History and Recent Developments

1.1 The computationalist background

The concepts of cognitive penetrability and impenetrability were originally deployed in the framework of a computationalist theory of mind by Zenon Pylyshyn (1980; 1984) and Jerry Fodor (1983). Pylyshyn and Fodor used the concepts to define the limits of a computationalist account of how the brain represents and manipulates the contents of propositional attitudes. In that framework, cognitively penetrable states and processes are those sensitive to the information involved in computations that perform thought and reasoning. Cognitively impenetrable processes either involve different computations, which are not sensitive to those that perform thought, or else are outside the scope of the computationalist approach to mind.

As part of the hypothesis that perception is cognitively impenetrable, Pylyshyn (1999; 2003) sought to isolate a set of early perceptual processes from thought and to describe them as distinct items in a hierarchical cognitive architecture. This was the first systematic empirical hypothesis to the effect that perception is cognitively impenetrable. Specifically, it claimed that part of visual perception was cognitively impenetrable. Pylyshyn calls this part of vision ‘early vision’. It includes the set of visual processes from stimulus onset up to the construction of viewer-centred (egocentric) volumetric representations of object surfaces. At this stage of vision, according to Pylyshyn, the scene should also be segmented into primitive visual objects, usually called ‘proto-objects’ (see Pylyshyn and Storm 1988; Pylyshyn 2007). Simultaneously, Pylyshyn offered a precisely formulated hypothesis about the penetrability of visual perception: he held that the remainder of vision—‘late vision’, which includes object recognition and identification—is performed jointly with contributions from long-term memories, semantic information (such as conceptually encoded information about kinds), agentively focused attention, and even conscious hypothesizing by the perceiving subject. For more information about Pylyshyn’s accounts of early and late vision, and a comparison with David Marr’s account of vision, see Section 2.2.

Fodor’s hypothesis that perception is modular is closely related to Pylyshyn’s hypothesis that perception is cognitively impenetrable. Fodor (1983: 73–83) uses the concept of cognitive penetrability to describe the kinds of influences on perception that the modularity hypothesis denies, and Pylyshyn (2003: ch. 2) uses Fodor’s concept of modularity to describe impenetrable visual processes. Fodor’s modules are information-processing brain mechanisms that perform circumscribed, partial tasks on limited and exclusive ranges of inputs. The role of perceptual modules is to yield outputs in a suitable format for use by cognitive centres by drawing only on perceptual inputs and perceptual processing resources. For example, Fodor favours a model for vision in which ‘the final stage of visual input analysis involves accessing a “form-concept” dictionary which, in effect, pairs 3D sketches with basic categories’ (1983: 137). On such a model, brain representations of three-dimensional objects are
generated from retinal stimulations by vision alone, without any help from memories or semantic information (such as conceptually encoded information associating generic shapes with kinds).

Fodor enumerates a number of properties to describe modular systems, but a system cannot be modular unless it is informationally encapsulated (Fodor 2001: 63; 1983: 71). For a computational system to be encapsulated, its informational resources have to be restricted to its proprietary database of information. This amounts to an absence of informational exchange between systems: ‘X is informationally encapsulated from Y if and only if X cannot use information from Y in its computations’ (Mahon and Wu, Chapter 8 of this volume).

If a perceptual information-processing system is modular in Fodor’s sense, it will be informationally encapsulated from other systems, including cognitive ones. It will not be sensitive to the information involved in the computations that perform thought and reasoning, and will thus be cognitively impenetrable. However, the cognitive impenetrability of a perceptual module does not imply its informational encapsulation. On the differences between cognitive impenetrability and perceptual modularity, see section 2.1 on encapsulation, and section 2.2 on early vision.

1.2 Cognitive impenetrability and New Look psychology

In addition to their positive role in computationalist accounts of cognitive architecture, the concepts of impenetrability and modularity were enlisted for criticism of competing views of the mind based on experimental findings in the ‘new look in perception’ movement in psychology. Those findings were interpreted as evidence that ‘all perceptual experience is necessarily the end product of a categorization process’, that ‘perception is a process of categorization in which organisms move inferentially from cues to category identity’ (Bruner and Goodman 1947), and that ‘the perceptual effect of a stimulus is necessarily dependent upon the set or expectancy of the organism’ (Bruner and Postman 1949, cited by Mole, Chapter 9). Pylyshyn, in turn, held that the view of the mind offered by Bruner and Goodman meant that ‘values and needs determine how we perceive the world, down to the lowest levels of the visual system’ (2003: 52; our emphasis). A programmatic statement of the New Look movement and its views on perception can be found in Bruner and Goodman (1947). For detailed overviews, see Gregory (1974; 1970) and Rock (1983). In this volume, see the contributions by Machery and Mole (Chapters 1 and 9).

A classic illustration of the relevant findings is provided by Bruner and Postman (1949) (for detailed discussion see Mole, Chapter 9). In Bruner and Postman’s experiment, when subjects who expected to see items belonging to certain categories (playing cards) were presented with the objects at sufficiently short intervals, they reported that the objects had the expected colours or a compromise between the expected and actual colour, even when the objects had anomalous colours (e.g. a red six of spades). Bruner and Postman explained the reports by claiming that expectations influence perception of the stimulus. If the expectations were determined by conceptual memory colours,
then conceptually encoded information or beliefs causally influence the mechanisms that generate perceived colour. Comparable findings on colour perception were reported by Delk and Fillenbaum (1965), and recently by Hansen et al. (2006) and Olkkonen et al. (2008); Hansen et al. concluded that their ‘results show a high-level cognitive effect on low-level perceptual mechanisms’. There are many similar findings in other areas such as auditory recognition of sentences, visual recognition of words, or dependence of size perception on value. If the interpretations are correct, the findings meet a condition for cognitive penetration of early vision given by Pylyshyn: ‘if a system is cognitively penetrable then the function it computes is sensitive, in a semantically coherent way, to the organism’s goals and beliefs, i.e., it can be altered in a way that bears some logical relation to what the person knows’ (1999: 343).

Pylyshyn (1999; 2003) used the CI hypothesis to argue against such interpretations of the experiments. He presented clinical findings (such as Humphreys and Riddoch’s 1987 work on visual agnosia) which suggested the dissociation of part of visual perception from cognition. He adduced psychophysical evidence (Sperling and Melchner 1978; Tanner and Swets 1954) that early vision resolves ambiguities in the scene autonomously from rational inference, arguing that many putative cases of cognitive influence on visual interpretation of stimuli were in fact hard-wired visual constraints impenetrable to cognitive states. Finally, he sought to show that ‘many apparent examples of cognitive effects in vision arise either from a postperceptual decision process or from a preperceptual attention-allocation process’ (Pylyshyn 2003: 90).

1.3 Cognitive architecture and epistemology

A third strand of the original cognitive penetrability debate, on which Fodor in particular focused, are the consequences of hypotheses about cognitive architecture for epistemology. New Look accounts of the nature of perception were used by philosophers of science to give a psychological defence of the idea that scientific observation is theory-laden (Hanson 1958; Kuhn 1962; Feyerabend 1962). (See Mole, Chapter 9: Kuhn used Brauer and Postman’s 1949 study, while Hanson drew on the phenomenology of perceiving ambiguous figures, which was popular at the time thanks to work by Gestalt psychologists and to Wittgenstein’s influence.) Duhem (1914) had argued that observations are framework-relative, and that observational data are never uninterpreted, theory-neutral descriptions of events; to this, Hanson, Kuhn, and Feyerabend added that what we see is already an interpretation of incoming information based on the perceiver’s theoretical commitments and conceptual frameworks—in other words, that perception is cognitively penetrated. This allowed them to challenge the view that scientific theories can be empirically tested and compared to other theories, since it left no theory-neutral ground on which to make a rational choice among alternative theories based on experimental outcome, making scientific theories potentially incommensurable. This outcome has been defended by constructivists in the philosophy of science who deny that scientific theories relate
observers to mind-independent objects (see Kukla 2000; Kitcher 2001). On the other 
hand, Brown (1987), Brewer and Lambert (1993; 2001), and Kitcher (2001) hold that 
perception is theory-laden, but simultaneously seek to alleviate the consequences of 
conceptual relativism by arguing that strong bottom-up components of perception 
provide empirical constraints sufficient to prevent constructivism.

Fodor dealt directly with the epistemological consequences of choosing a theory of 
cognitive architecture. Paul Churchland (1979) had argued that observation is theory-
laden by relying on new findings about the plasticity of perception from cognitive 
neuroscience, connectionism, and vision science. In seeking to rebut Churchland’s 
claims, Fodor explicitly stated what he took to be the epistemological relevance of the 
impenetrability and encapsulation hypotheses: ‘The question at issue is: What are the 
psychological conditions under which differences among the theories that observers 
hold are not impediments to perceptual consensus among observers. Cognitive encap-
sulation seems to be an empirically necessary condition for this’ (Fodor 1988: 189). 
Like Pylyshyn, Fodor used the thesis (inspired by Marr 1982; see section 2.3) that 
putative instances of theory-ladenness were in fact hard-wired modular constraints 
on visual processing: perceptual modules apply general assumptions inherent to their 
functional architecture to transform proximal stimuli into brain representations which 
can be exploited cognitively. For observation to be theory-laden, Fodor pointed out, 
perceptual processing would have to access the contents of propositional attitudes, and 
this appears to be refuted by the impermeability of perceptual illusions to the contents 
of beliefs. Two points made by Churchland (1988) in reply to Fodor were to prove 
particularly resilient: (a) that the existence of abundant top-down neural pathways 
from higher cognitive centres to the circuits of low-level vision can be explained only 
if one assumes that those pathways allow a transfer of information from cognitive areas 
in the brain to perceptual processing sites; (b) that perceptual plasticity of the brain in 
general, and the areas dedicated to perceptual processing in particular, refutes Fodor’s 
claim that perception is performed by encapsulated modules.

1.4 Recent developments

The hypothesis that early vision is cognitively impenetrable has been defended exten-
review of recent accounts of attention, arguing that they confirm Pylyshyn’s thesis that 
attentional influence on perception is restricted to late vision and the pre-perceptual 
allocation of attention. He also presented an analysis of recent neuroscientific evidence 
about the latencies at which visual processes occur (esp. Lamme 2000; 2003; 2005; 
Lamme and Roelfsema 2000) in order to determine whether the outputs of early vision 
are representations of which subjects can be phenomenally aware and can 
be reported by subjects (see Block 2007a; 2007b)—a point that needs to be settled 
if the outputs of putatively impenetrable states are to be epistemically useful. An 
account of the relations between cognitive impenetrability and the hypothesis that 
perception has nonconceptual content was offered by Raftopoulos and Muller (2006)
and Raftopoulos (2009; 2014), who claimed that the CP of early vision is a necessary condition for perception to have nonconceptual content. This position was disputed by Bermúdez and Cahen (2011) and is discussed in Macpherson’s contribution to this volume (Chapter 13). Finally, Raftopoulos (2001a; 2001b; 2006) has used the CI hypothesis to argue against theory-ladenness and in favour of scientific realism. (On the role of different forms of attention, see Section 2.6; on Lamme and Roelfsema’s work, see 2.2; on awareness and consciousness, see 4.4; on nonconceptual content, see 4.3; on the theory-ladenness of observation, see 2.5.)

Siegel (2006; 2010; 2012; 2013) has shown that questions about cognitive penetrability, and its potential impact on justification, are entangled with the questions of whether kinds are represented in perception and whether it is possible to distinguish visual from cognitive phenomenology. Siegel’s (2006–2010) thesis that visual experience is affected by concept possession and that the contents of perception include high-level properties such as kinds implies a form of cognitive penetration of perception, although it is not clear whether early or late perceptual processes would be affected. Siegel (2012) describes ‘illicit feedback loops’ from cognitive to perceptual states which have the potential to vitiate the perceptual justification of beliefs: if a perceptual experience has been formed through processes affected by beliefs, and the affecting beliefs are related to the belief that is being evaluated, justification of that belief by the perceptual experience will be epistemically suspect. (See Section 4.1 on CP and the contents of perception, and 4.7 on CP and perceptual justification of belief.)

Lyons (2001; 2009; 2011; and Chapter 3 in this volume) has made substantive proposals about the nature of perceptual modularity and studied the epistemological consequences of cognitive penetrability and perceptual unencapsulation. On the moderate conception of modularity that he puts forward, a system can be modular even if it is not informationally encapsulated (see Section 2.1 on modularity; Lyons, Chapter 3; and Lyons 2001). Lyons denies that cognitive penetration of perception would necessarily vitiate perceptual warrant, arguing that some forms of cognitive penetration are epistemically benign and others vicious, depending on whether penetration increases or diminishes the reliability of perception. For instance, cognitive influence resulting from slow learning will tend to be triggered by ‘enduring, stable features of the world’ and thus to be more reliable, unlike top-down effects which can be generated as a result of recent and local knowledge. (On the consequences of CP for perceptual warrant, see 4.7.)

Experimental findings on colour (Delk and Fillenbaum 1965; Hansen et al. 2006; Levin and Banaji 2006; Perky 1910) were used by Macpherson (2012, and Chapter 13) to argue that colour experience is cognitively penetrable. On her account of the experiments, the way we categorize objects determines which colour experience we have of them. Macpherson specifically argues that it is the phenomenal looks of objects that are affected by memory colours (which are conceptually encoded states), not our beliefs or judgments about the colours of the objects. She also describes a mechanism
which could allow the phenomenal character of nonperceptual, doxastic states, such as imagining, to interact with the phenomenal character of perceptual experiences. Rebuttals of the claim that colour perception is cognitively penetrable argue that the experimental findings can be explained at the level of perceptual judgment (Lyons 2011; Zeimbekis 2013) or appeal to the presence of early visual shape–colour associations which do not entail the influence of conceptually encoded memory traces (Deroy 2013; Lyons 2011; Raftopoulos forthcoming).

Much work currently under way on cognitive penetrability asks how it would affect the outcomes of existing philosophical debates, especially concerning perceptual warrant, nonconceptual content, perceptual and cognitive phenomenology, forms of awareness and consciousness, the role of attention in perception, realism and representationalism, and action. We explain the relevance of the penetrability debate to each of those topics in section 4. Before doing so, we have to describe some of the central concepts proper to the penetrability debate (Section 2) and review existing definitions of CP (Section 3).

We end this historical outline with a methodological consideration. While the concepts of cognitive penetrability and impenetrability were originally introduced in order to defend the impenetrability of perception in the framework of a computationalist theory of the mind, much recent work on the topic tends to argue in favour of the penetrability of perception. It does so either by citing experimental findings and phenomenological considerations, or by loosening the definition of penetrability to make it more inclusive (see Section 3). Skepticism about this reversal is expressed in the contribution by Edouard Machery (Chapter 1, ‘Cognitive Penetrability: A No-Progress Report’). While Machery does not make any substantive claims in favour of impenetrability, he seeks to ‘rekindle the doubts against the cognitive penetrability hypothesis’ on methodological grounds. He draws a historical comparison between the problems afflicting original cognitive penetrability claims, based on New Look psychology in the 1960s and 1970s, and shortcomings in the interpretations of experimental data offered by recent defenses of cognitive penetrability. Among the recurrent flaws pointed out in the findings and their interpretations is an inability to determine the locus of cognitive penetration. For example, in Delk and Fillenbaum’s (1965) colour experiments, there was no way to control whether it was perceptual experience or judgment that was affected by colour memories; but only an influence of memory on perception would count as CP. Another key set of recent findings for the penetrability debate, from Balcetis and Dunning’s (2007) studies of distance perception, is based on estimates that subjects made from memory, so it is not clear whether the subjects reported their perceptions or their memories. Yet Balcetis and Dunning conclude from their findings that ‘the impact of motivational states extends from social judgment down into perceptual processes’. In fact, Machery points out that New Look psychologists consciously dismissed the problem of knowing whether it is perceptual experiences or judgments that are influenced by cognitive states (‘it is not particularly important to find out whether the stimuli are really “seen” as larger or smaller’—Tajfel
Another recurrent problem, according to Machery, is failure to replicate key experiments. One of the most paradigmatic New Look studies, Bruner and Goodman’s (1947) experiment on how value affects size perception, failed to be replicated on several attempts (in 1949, 1966, and 2003); the same applies to recent experiments by Proffitt et al. (2003) on distance estimates. The replication problem is compounded by the protocols used in several studies, which did not guarantee normal viewing conditions or involved stimuli that were ambiguous and degraded. Machery concludes that recent empirical research and theoretical arguments in favour of cognitive penetrability ‘suffer from the exact same kind of problems that in the first place resulted in widespread doubts’ about the findings of New Look psychology and their interpretation.

2 Key Concepts of the Cognitive Penetrability Debate

The concepts of modularity, informational encapsulation, early vision, late vision, hard-wired visual processing, theory-ladenness, perceptual learning, different concepts of attention, and the perception/cognition distinction itself, are constitutive of the cognitive penetrability debate. In this section, we explain those concepts and distinctions and their roles in the debate.

2.1 Cognitive impenetrability, modularity, and informational encapsulation

As explained in Section 1.1, if a perceptual system is modular in Fodor’s sense, it will be cognitively impenetrable, but the cognitive impenetrability of a perceptual module does not imply its informational encapsulation. A perceptual information-processing system can be impermeable to cognitive influence but open to influence from other perceptual modules, including those that process information from other sensory modalities. Pylyshyn’s hypothesis that early vision is cognitively impenetrable is consistent with the occurrence of such interactions between different sensory modalities during perception, and this makes it more resilient than Fodor’s hypothesis about perceptual encapsulation. An example of crossmodal interaction is the McGurke effect (McGurk and MacDonald 1976): /ba/sounds paired with lip movements appropriate to the utterance of /ga/ result in perceptions of the sound as /da/, suggesting that auditory processing receives inputs from vision. If information from sensory modality A affects which outputs sensory modality B generates, then B is not informationally encapsulated from A and is not modular in Fodor’s sense: its information-processing resources are not restricted to its proprietary database of information. However, as long as B is not influenced by a cognitive process, it is not cognitively penetrable.

Machery also considers, and rejects, Pylyshyn’s (1999) proposal for resolving the locus problem by using signal detection theory. Pylyshyn sought to pin findings about penetration on changes in subjects’ response bias parameters, which would mean that the locus of penetration was judgment, not perception.
Daniel Burnston and Jonathan Cohen (Chapter 4, 'Perceptual Integration, Modularity, and Cognitive Penetration') outline a concept of modularity which is compatible with evidence that perceptual processes are unencapsulated with respect to one another. They outline an integrative model of perception designed to deal with interactions across and within sensory systems; then, instead of rejecting modularity, they replace the criterion of informational encapsulation with anisotropy. Fixation of belief is a characteristic example of an isotropic process: the holistic nature of the abductive inferences that the brain uses (Fodor 2001) means that information stored anywhere in the system may affect the formation of a belief. According to Burnston and Cohen, anisotropic systems are modular ‘just because, and in so far as, there is a delimited range of parameters to which their processing is sensitive’. On this account, if photoreceptors turn out not to be informationally encapsulated (see Broackes 2009), they will nevertheless be modular, while systems in charge of representing certain higher-order, multimodal spatial relations (Koechlin et al. 2003) could turn out to be non-modular. On the other hand, processes that are standardly thought of as cognitive—such as recognizing scenes of chasing (Gao and Scholl 2011) and certain social cues (Langton et al. 2000)—might turn out to be modular if they rely on ‘a significantly delimitable range of input parameters’. (Such cases of recognition are discussed again in Section 2.7.) Apart from modularity, Burnston and Cohen argue that a perception/cognition distinction, and the cognitive impenetrability of certain perceptual systems, can also be preserved in an integrativist framework.

Lyons (Chapter 3) proposes a concept of modularity based on the criteria of isolability, unitariness, and specialization to information-processing systems. The key difference between his definition of modules and Fodor’s can be grasped through Lyons’s criterion of isolability, which is weaker than informational encapsulation. For information-processing systems S1 and S2 and task T, ‘S1 is isolable from S2 with respect to T’ if S1 computes T and could do so even if S2 didn’t compute any functions’. On this definition, S1 is an unencapsulated system (it can draw on S2’s information-processing resources), yet it is modular. Lyons argues that this definition of perceptual modules can capture much of what Fodor required of informational encapsulation when it comes to safeguarding perception from epistemically harmful cognitive influences.

Both Lyons’s and Burnston and Cohen’s weakening of the criteria for modularity can be used to reject Fodor’s encapsulation hypothesis in favour of the hypothesis that early vision is cognitively impenetrable. This move also clearly limits the scope of criticisms of perceptual modularity (such as Prinz 2006) to encapsulation, as distinct from cognitive impenetrability.

Ophelia Deroy (Chapter 5, 'Multisensory Perception and Cognitive Penetration: The Unity Assumption, Thirty Years After') studies interactions between sensory modalities that seem not to be restricted to low-level perceptual processing, but to be mediated by cognitive information. Deroy points out that, while input modules are usually seen to ‘all fit at the same place in the hierarchy’ between sensory systems
and systems that implement thought-like processes, there may also be multisensory modules—for instance, visual and auditory spatial information may 'feed into a multisensory system which generates representations of audio-visual localization'. If such modules were appropriately influenced by semantic information, perception would be cognitively penetrated. That appears to be the case when multisensory interaction exhibits the phenomenon of 'semantic congruence', in which background assumptions about particular objects or combinations of sensory features determine whether or not information from different sensory sources will be integrated into a single object. For example, kettle shapes and whistling sounds are integrated into a single audiovisual object, while kettle shapes and barking sounds are not (Jackson 1953). Such assumptions apparently involve stored conceptual representations of objects, suggesting that their influence amounts to cognitive penetration. Deroy reconstructs and assesses the empirical reasoning underlying the interpretation of congruence as a semantic effect, concluding that the evidence does not show congruence to be a top-down causal influence. She also argues that if a top-down causal explanation of congruence turned out to be preferable, it would still need to be supplemented by an argument showing that the top-down influences are cognitive ones specifically. An alternative explanation of the evidence is that it results from 'crossmodal correspondences': non-rational representations of congruence which do not involve conceptual representations, such as the documented association between high pitches and bright surfaces (Spence 2011).

2.2 Cognitive penetrability, early vision, and late vision

The distinction between early vision and late vision became decisive for the cognitive penetrability debate when Pylyshyn used it to isolate the part of perception that he considers cognitively impenetrable. Note that the early/late vision distinction does not coincide with the distinction between cognition and perception, which is the subject of Section 2.7.

Fodor, as we saw earlier (14), holds that the informationally encapsulated part of vision encompasses all of the processes that lead up to visual object recognition. But that is unlikely to be the case; it is widely thought that object recognition is secured by an interaction of visual information with conceptual information and long-term memories (see Peterson 2005 for a review of object-recognition hypotheses). In keeping with this, Pylyshyn (1999; 2003) isolated an early stage of vision that excludes recognitional processes and extends no further than the visual representations that Marr (1982) calls the $2\frac{1}{2}$D sketch. He claimed that only this early stage of vision is cognitively impenetrable.

The distinction between early and late vision is central to Marr's account of vision. According to Marr, vision first extracts information about the edges of objects in a scene, mainly from the way the scene reflects light (1982: ch. 2). From that information jointly with further information about lighting, and with processes that include stereopsis, parallax, the hard-wired interpretation of monocular depth cues, and
geometric and topological principles for generating volumetric shape representations, the visual system constructs a viewer-centred (egocentric) representation of volumes and depth relations in the scene (Marr 1982: chs 3 and 4). Marr calls this type of visual representation the $2\frac{1}{2}$ D sketch; it is no less volumetric than a 3D sketch, but being a spatial representation of a three-dimensional scene from a limited viewpoint, without additional semantic interpretation, it is restricted to visible surfaces. The invisible parts of objects are only represented subsequently by what Marr calls the 3D model, which is suitably formatted for matching with memorized object representations.

Marr uses the terms 'early vision', and sometimes 'pure perception', to describe the kinds of processes that go into constructing the $2\frac{1}{2}$ D visual sketch: 'the recovery of surface information by purely data-driven processes without the need for particular hypotheses about the nature, use, or function of the objects being viewed' (Marr 1982: 269). He thus sees vision up to that point as involving no processing that we could call semantic.

Among the evidence used by Marr to support this position were findings in clinical neurology. Warrington and Taylor (1973; 1978) found that right parietal lesions caused difficulty in recognizing objects only when they were viewed from angles which foreshortened their natural axes (a geometric property exploited by early vision to construct depth from contours). Marr (1982: 328) saw this as evidence that shape is constructed by the visual system independently of semantic input about the identity of objects (information about what kinds of objects they are). Generally, his position was that viewer-centred volumetric shape can be determined by vision alone even in difficult circumstances, in which one may have expected that vision draws on semantic information to supplement or disambiguate visual output (1982: 36). Marr contrasted his account of vision with models given by computer vision scientists (e.g. Freuder 1983), whose object-recognition programs used semantic information to determine shapes and segment objects in a scene. Such models suggested that visual object recognition is determined throughout by semantic information. (Similar criticism of models provided by computer vision scientists can be found in Pylyshyn; see Pylyshyn 2003: 62.) Instead, Marr thought that semantic information only intervenes to assist object recognition at a later stage of vision which presupposes the $2\frac{1}{2}$ D sketch (because to reach that stage, vision first has to map the $2\frac{1}{2}$ D sketch’s viewer-centred coordinate frame to an object-centric coordinate system; see Marr 1982: 299–302).

To distinguish early vision from late vision, Pylyshyn (1999: 2003) drew on clinical evidence of functional dissociations of visual and cognitive functions; on the persistence of visual illusions; on the independence of principles of visual organization from principles of inference (e.g. in Kanizsa’s visual completions); and on psychophysical evidence of the attenuation and gating of visual signals. Humphreys and Riddoch’s (1987) work on visual agnosia has shown that there can be deficits in object recognition despite the preservation of both memory and visual computations. In their words, this supports the view that “perceptual” and “recognition” processes are separable; and that the perceptual representation used in recognition can be “driven” solely by
stimulus information, so that it is unaffected by contextual knowledge’ (Humphreys and Riddoch 1987: 104; cited by Pylyshyn 2003: 72). Pylyshyn also discusses the dissociation between recognition and visually guided action (Milner and Goodale 1995; Weiskrantz 1995; Holmes 1998), concluding that the outputs of the dorsal stream are not ‘available to the rest of the mind/brain’ (2003: 154). (For an explanation of what the dorsal and ventral visual streams are, see Mahon and Wu, Chapter 8, §3.)

Pylyshyn’s description of the limits of early vision is fundamentally similar to Marr’s because it is based on Marr’s hierarchic conception of vision. Early vision is described as extending no further than viewer-centred representations, and the feedback of semantic information to early visual processes is admitted only after early vision has already assigned shape. The outputs of early vision are described as consisting of ‘shape representations involving at least surface layouts, occluding edges (where these are parsed into objects), and other details sufficiently rich to allow looking up parts of the stimulus in a shape-indexed memory for identification’ (2003: 91). Pylyshyn adds two key features to Marr’s account. One is the thesis that early vision constructs primitive object-like representations (Pylyshyn 2007; Pylyshyn and Storm 1988), as distinct from generating a viewer-centred representation of the scene without differentiation into objects. The second is the thesis that attention does not directly modulate early visual processing; we discuss this in Section 2.6 on attention and cognitive penetration.

Recent neuroscientific work on the timing of neural processes in the visual system allows us to form a working hypothesis for defining early vision in terms of processing latencies (Lamme 2000; 2003; 2005; Lamme and Roelfsema 2000; Roelfsema et al. 2000; Roelfsema 2005). That research suggests the following three-stage picture of vision.

(1) Signal transmission from the retina through the visual areas of the brain up to the inferior temporal cortex lasts for about 100 ms after stimulus onset. This flow of information is described by Lamme and Roelfsema as a ‘feedforward sweep’ because it is a form of signal transmission immune to feedback from areas further downstream; the reason for this is that neurons in each area fire at a much slower frequency than neurons in the earlier area, ‘leaving no time for lateral connections and no time for feedback connections to exert their effect’ (Lamme and Roelfsema 2000: 572). Thus, for as long as the feedforward sweep of information through the visual areas lasts, it is a bottom-up form of signal transmission. The feedforward sweep is unconscious; it determines the classical receptive field of neurons and their basic tuning properties, extracts information which is subsequently useful for categorization, and results in some initial feature detection.

(2) Feedback (‘recurrent’) connections and lateral connections between neurons begin to take place at the end of the feedforward sweep, culminating at about 120 ms; information is fed back to earlier areas (such as VI; see Lamme and Roelfsema 2000), resulting in interaction between information distributed along the visual stream. At this stage, the visual system forms representations of transducible features (spatiotemporal properties, surface properties, viewer-centred shape, colour, texture, orientation,
motion, etc.) and performs primitive feature binding and object segregation. This form of feedback is not thought to involve signals from cognitive centres but is considered to be confined to interactions within the visual areas of the brain; this is why Lamme and Roelfsema call it 'local recurrent processing'.

(3) Finally, at 150–200 ms, signals from frontal and prefrontal areas and mnemonic circuits begin to intervene and modulate perceptual processing in the visual cortex; eventually, the recurrent interactions with areas outside the visual stream make storage in visual working memory possible and give rise to global recurrent processing. (See Lamme 2003, 2005. These findings are supported by the 'global neuronal workspace model': Dehaene et al. 1998; Dehaene et al. 2003; Dehaene and Changeux 2005; Sergent et al. 2005). Visual processing, which up to this point comprised bottom-up processes, lateral connections, and local top-down effects from one visual area to another, can now be modulated by the activation of information stored in long-term memory as the synaptic weighting of neurons. Thus, global recurrent processing potentially marks the onset of conceptualization of perceptual content; this stage of visual processing is conceptually modulated and cognitively penetrated.

Raftopoulos (2009) holds that the feedforward sweep and local recurrent processing jointly form a suitable candidate for the role assigned to early vision by Pylyshyn, and that global recurrent processing in visual areas corresponds to late vision. For early vision to be cognitively impenetrable, either it has to receive no top-down signals from cognitive centres, or, if it does, the signals must not affect it in a 'semantically coherent way' or in a way that alters the content of early vision (see Section 3 on definitions of cognitive penetrability). This would have to apply equally during global recurrent processing. According to Raftopoulos, the feedforward sweep and local recurrent processing are impenetrable in this sense. Raftopoulos also argues that Lamme and Roelfsema's distinction between local and global recurrent processing is consistent with Pylyshyn's account of the outputs of early vision; and that, in keeping with Pylyshyn's account, visual processing during the feedforward sweep and local recurrent processing are not directly influenced by cognitively driven attention.

2.3 Cognitive penetrability and hard-wired ‘perceptual knowledge’

A top-down flow of information is the transmission of signals from a later stage of an information processing stream in the brain back to an earlier stage. Churchland describes such effects when, citing evidence by Wolter and Lund (1968), he writes that 'there is some evidence that fully 10 percent of the axonal fibers in the human optic nerve are descending projections from the LGN [lateral geniculate nucleus] back to

---

4 Contextual or horizontal modulation and its effects are also well documented; see Lamme 1995; Lamme and Spekreijse 2000; and Raftopoulos 2009: 34–5, for further references. There is ample evidence that the early vision module consists of a set of interconnected processes for orientation, shape, colour, motion, stereo, and luminance that cooperate within it (Cavanagh 1988; Zeki 1981; 1993; and Raftopoulos 2009: 35, for further references).
the retinal surface’ (Churchland 1988: 177). This evidence has been confirmed by Van Essen (1985) and more controversially by Zeki (1993). Generally, local feedback signals are thought to be essential to vision from an early stage (see also Zeki 1981; Cavanagh 1988; Lamme and Spekreijse 2000).

Evidence of top-down effects on vision is often used to argue that vision is cognitively penetrable, and this is how Churchland was using it in the passage cited. However, for a top-down effect to amount to cognitive penetration of perception, a perceptual process has to be affected by another process downstream in the cognitive areas of the brain. For example, the fact that there are top-down effects on area V1 from V4 does not mean that there is cognitive penetration of vision, because V4 is a visual area which does not contain any information we could plausibly count as cognitive, such as long-term memories, concepts, or the contents of propositional attitudes.

Many such forms of local top-down visual processing can be mistaken for cognitive influences on early vision, because the computations they perform are consistent with assumptions about the physical world. Vision has evolved to function according to those assumptions and they are built into the neural circuits that subserve visual processing. For example, Marr shows that the design of physiological visual mechanisms implements Bayesian priors to transform light differentials into edges, calculate depth from stereopsis or curvature from smooth variations in illumination, or implement the epipolar constraint. He describes the assumptions as ‘hard-wired’ into the visual system (1982: 185) and as reflecting ‘some kind of a statistical rule of the universe’.5 According to Marr and other vision scientists, the assumptions are necessary because the retinal image under-determines both the distal object and the percept; unless vision was constrained by them, it would not construct perceptual representations of distal objects. Visual processing is constrained by such principles, which modulate information processing, at every level. Cavanagh (2011: 1545) holds that vision does not rely on purely bottom-up analyses drawing on retinal information, but instead completes construction of the percept by drawing on ‘object knowledge’, a set of rules that guide or constrain visual processing in order to resolve underdetermination. Rule-based extensions from partial data would constitute a form of inference. Spelke (1990) holds that perception is constrained by a number of domain-specific principles about material objects and some of their properties. These constraints involve attentional bias toward particular inputs and a certain number of principled predispositions constraining the computation of those inputs; their operation is described by Spelke as inferential and akin to thinking.

5 Perhaps because Fodor (1983) uses the expression ‘hard-wired’ to describe modular processing, the expression has become almost synonymous with ‘modular’ and would thus appear to imply cognitive impenetrability. But we are not making the point that principles like those Marr describes for visual processing are modular and thus impenetrable; our point is that knowledge of the principles is irrelevant to their application to stimuli by vision.
However, such 'hard-wired' principles, 'formation principles' (Burge 2010), or 'operational constraints' (Raftopoulos 2009) are not available to introspection; they function outside the realm of consciousness and their operations cannot be attributed as acts to perceivers. When perception parses objects according to the Spelke’s principles that objects move in continuous paths and are rigid, subjects do not have to believe those principles; instead, they constrain perception’s modus operandi. Unlike theoretical commitments, they are only used by perception and are not available for a wide range of cognitive tasks. The constraints cannot be overridden since they are not under the perceiver’s control; subjects cannot substitute them with another body of constraints even when they know that they lead to errors.

One way to account for the status of such apparent assumptions in perception is to hold that while perception operates in accordance with them, they are not represented even at a sub-personal level. They could be design-level properties that do not require the organism to token any representations of rules (see Dennett 1971). In that case, a neural state is formed through the spreading of activation and its modification as it passes through synapses; hard-wired constraints are implemented by computational processors, and computational principles describe the transformations from one state to another. Although the states produced by transformations have contents, the mathematical principles are not states of the system and do not have to be represented by it. This view has been defended by Burge:

For many philosophers, the notion of computational states or explanations is theory-laden in a way that I do not intend. When I call states or explanations 'computational', I do not mean that there are transformations on syntactical items, whose syntactical or formal natures are independent of representational content [of the computed states]. I also do not mean that the principles governing transformation are instantiated in the psychology, or 'looked up', even implicitly in the system . . . principles governing perceptual transformations . . . are not the representational content of any states in the system, however unconscious. (Burge 2010: 95)

Alternatively, the rules or generalizations could be realized by computational processors in the brain as a form of ‘tacit knowledge’ not realized by attitude states. Davies (1989) holds that such knowledge is not inferentially integrated with attitude states but is subdoxastic and exists in special-purpose, separate subsystems (see also Stich 1978). The constituent concepts of attitude states such as beliefs must be concepts possessed by the believer, while the contents of tacit states are not conceptualized and not even accessible to the subject. Similarly, Stich and Nichols (1992: 46) describe a ‘tacit theory’ as one which is not sententially encoded.

In neither case would the principles reflected by operational constraints on perception be rules of inference that the visual system looks up to perform transformations, or premises used in such inferences. Not only are operational constraints not conceptually and propositionally structured mental contents, but they do not even seem to be contentful states of the perceptual system. Finally, the fact that perception relies on such constraints does not entail that perception is affected by concepts, since
on any interpretation of the principles implemented by operational constraints, those principles are not conceptually encoded.

2.4 Cognitive penetrability and perceptual learning

Even when one perceptual information-processing system draws on another perceptual system’s resources in a way that is infra-perceptual and clearly non-cognitive, it is still possible to argue that its outcome is the same as that of cognitive penetration. One such form of influence, which occurs without any contribution from personal-level or doxastic states, is the effect of perceptual learning on perceptual processing, in which the subject’s past experiences partly determine how her visual states process incoming information.

There is evidence that visual memories acquired in the course of perceptual learning affect the way we perceive the world. The perceptual tasks affected are those that psychologists call object classification, identification, and categorization. The uses of the terms in psychology do not coincide with their uses in philosophy (see Section 4.5, and Lowe, Chapter 14). The task that psychologists call object classification is a coarse-grained form of classification which occurs at very short latencies: 95–100 ms and 85–95 ms after stimulus onset for classifying items as animals and faces respectively. Object identification and categorization are lengthy processes; they start about 150 ms after stimulus onset and last until 300 and even 600 ms (Johnson and Olshausen 2005). When the effects of familiarity on visual processing occur at 300–360 ms, which they often do, they are post-sensory, affecting the semantic information and processing required for identification and categorization (Delorme et al. 2004). As such they could be used to claim the cognitive penetrability only of late vision.

However, Liu et al. (2009) and Peterson and Enns (2005) have shown that familiarity with objects or scenes built through repeated exposure, and sometimes through a single presentation, facilitates visual search, affects figure–ground segmentation, and speeds up object classification (in addition to identification). Crouzet et al. (2010) have also shown that object classification can be affected by familiarity and repetition memory. These findings could suggest two potential sources of cognitive penetration.

The first is that, since classification occurs at very early latencies, such effects could not be considered post-sensory and would potentially penetrate early vision. This threat would materialize if the classification process required either semantic information to intervene, or the representations of objects in working memory to be activated, since that too would amount to conceptual involvement in the visual process. However, the effects of familiarity on early classification seem to result from low-level visual areas: from V1 to V4 (Kirchner and Thorpe 2006) and perhaps a bit more upstream from the posterior IT cortex (Peterson 2003) and lateral occipital complex (Grill-Spector et al. 1998). Besides, if the classifications affected by visual memories required semantic information or activation of object memories, they would not be as fast as they are. (For further studies confirming that the effects are
low-level, see Chaumon et al. 2008, Grill-Spector et al. 2006, Peterson 2003, Ullman et al. 2002, Delorme et al. 2004, and VanRullen and Thorpe 2001.) Ullman et al. (2002) show that statistical differences in physical properties of different subsets of images are detected very early by the visual system before any top-down semantic involvement, and suggest that early visual areas store implicit associations representing fragments of objects and shapes; according to Peterson, early visual memory stores ‘edge complexes’ as opposed to whole objects and object-shapes.

The second potential source of cognitive penetration is that even if visual memories affected early processing without the mediation of either concepts or personal-level states, they would still mean that our past experiences shape the way we see the world, constituting what Stokes (Chapter 2) calls ‘diachronic cognitive penetration.’ That what we see is already an interpretation of incoming information is also true of the hard-wired formation principles discussed in Section 2.3; but in those cases, what we see is a function not of past personal experiences but of the way the visual system evolved to reflect high-level generalities about the distal objects presented to perception. In the case of visual memories which are proper to early vision and affect its processing of incoming data, individuals with different experiences see the world differently.

Thus, although the findings about visual memory do not actually show influences on perception from cognitive states, they could be used to defeat epistemological objectives associated with denials of cognitive penetrability. As explained in Sections 1.2 and 1.3, the cognitive impenetrability and perceptual modularity hypotheses were designed to rebut epistemological views flowing from certain interpretations of psychological findings, so the indirect form of perceptual bias caused by perceptual learning would strike at the heart of both hypotheses. Stokes (Chapter 2) explicitly incorporates considerations about theory-neutrality and perceptual justification in his definition of cognitive impenetrability; and like Fodor (1983; 1988), Raftopoulos (2006; 2009) sees claims about cognitive architecture as inseparable from theses in the epistemology of perception. This takes us to our next topic, the relations between cognitive penetrability and theory-ladenness.

2.5 Cognitive penetrability and the theory-ladenness of observation

Bogen (2010) recently distinguished three forms of observational theory-ladenness: (i) theories affect perceptual processes so that the percept is partially determined by our theoretical frameworks; (ii) observations cannot be described in a theory-neutral way and the meaning of observational terms is determined by theoretical presuppositions; (iii) theories make certain observations more salient than others. The concept of theory-ladenness used in the CP literature is the first one, which affects perception. As we saw in section 1, Hanson (1958) and Kuhn (1962) supported their claims that observation is theory-laden by citing psychological findings which suggest that perception is penetrable, so they thought that penetrability makes observation theory-laden. Simultaneously, defenders of the impenetrability hypothesis deny that observation is theory-laden (Fodor 1988; Raftopoulos 2006; 2009). However, they
do not proceed by denying theory-ladenness in order to block the conclusion that perception is cognitively penetrable; instead, Fodor, Raftopoulos, and sometimes Pylyshyn (2003) start by trying to show that perception is impenetrable and that therefore observation is not theory-laden. A potential danger for that strategy lies in the fact that cognitive impenetrability may not suffice to keep theory-ladenness at bay: as Bogen’s other senses of the concept of theory-ladenness show, observation could be theory-laden in some other sense even if perception was cognitively impenetrable. A similar conclusion can be drawn from a concept of theory-ladenness recently proposed by Lyons (2011), on which the formation of perceptual beliefs is conceptually driven. If perception is taken to include perceptual belief, which is a doxastic state, then part of perception is theory-laden. But that part is unlikely to be an early part of perception; and if the CI hypothesis is restricted to early perceptual processes, CI is likely to be compatible with the form of theory-ladenness described by Lyons.

Whether observation is theory-laden, and whether it is theory-laden as a consequence of cognitive penetration of perception, depends on what it is for a subject to possess a theory and on whether the influence of the theory can amount to cognitive penetration. Any concept or belief which affects how perceptual experience is generated counts as a penetrating state, irrespective of whether it is part of a scientific theory, a folk theory, or not essential to any particular theory. However, having a theory does not always mean having a set of beliefs and concepts. Theories may be tacit as opposed to ‘internally represented knowledge structures that [invoke] explicit rules or explicit sentence-like principles’ (Stich and Nichols 1992). This is likely to apply to our folk physical theory about the behaviour of everyday objects, and may or may not apply to the brain’s diachronic adaptation to wearing inverted lenses (described by Churchland 1988). It is not known whether the effects, if any, of folk physics on perception would constitute cognitive penetration. Pagondiotis (Chapter 15; see Section 4.5) argues that visual experience is penetrated by practical non-propositional knowledge. It is sometimes claimed that folk psychology (construed as a theory) is modular, which would make it impenetrable to other beliefs (Jackendoff 1992). This position has been criticized on the plausible grounds that social reasoning does not draw on limited resources or even limited sets of beliefs (Currie and Sterelny 2006)—i.e. that social reasoning is isotropic. Finally, the concept of theory-ladenness should not be used to describe the fact that a perceptual system applies the kinds of formation principles or operational constraints described in Section 2.3. For the reasons pointed out there, while the functioning of perceptual systems can be described as if they were applying a theory, there is no theory, and perhaps there are even no representations, involved in those transformations; and in any case not a theory known or applied by subjects.

An interesting form of theory-ladenness is that which potentially stems from the influence of perceptual learning on perceptual processing. As seen in the section on perceptual learning (2.4), although the findings about visual memory do not show
influences on perception from cognitive states specifically, they could still be used to defeat epistemological uses of the cognitive impenetrability hypothesis. Suppose that through perceptual learning by means of repeated visual experiences in her field, a scientist's specific professional needs have shaped her perceptual sensitivity so that she can recognize patterns that others cannot. She has learned which dimensions of visual analysis to attend to, and this process has reshaped her basic sensors by selecting the output of certain feature detectors. Suppose that this learning has induced changes in the circuitry of her early vision altering her visual perception; and that her theoretical commitments allow her to focus on specific locations of the visual array in her field and synthesize the picture in a way that is unavailable to others who do not share her visual training and memories. It is true that scientists (e.g. radiologists) and other experts (e.g. birdwatchers) who are trained to discern certain patterns and have memorized them can perceive patterns that other subjects cannot. Similarly, the scientist can detect patterns that others cannot.

If this scenario is accurate, it has significant consequences for the role of theory-ladenness in the penetrability debate. It suggests that non-cognitive, clearly perceptual influences on incoming visual information can be indirect bearers of the kinds of theoretical commitments that we usually think of as the content of conceptually couched theories. The outcome would vindicate Churchland's belief that theory-ladenness of observation flows from the brain's plasticity. It would also refute the implicit assumption, described above, that the cognitive impenetrability of perception prevents theory-ladenness: here, the influences on vision are not cognitive, so impenetrability can be upheld; yet perception ends up being theory-laden.

Since what is at stake are the epistemological consequences of discoveries about perceptual processing, it is worth asking whether the scenario described, if it was realistic, would affect scientists' assessments of experiential evidence in the context of theory evaluation. One would expect perceptual learning, which is a honing or improvement of perceptual discernment in a specific domain, to be a benign influence on perception and not to lead to incommensurability of the learner's percepts with those of other perceivers. Yet Kuhn argued that theories are incommensurable precisely on the grounds that observers presented with the same stimuli can see different things if their brain circuits are shaped differently by perceptual learning. Perhaps one could argue against Kuhn that, since both slow and fast perceptual learning are data-driven and not cognitively driven, if two subjects developed perceptual skills under different Kuhnian paradigms but were trained with the same data, they would end up seeing the same things. In that case, perceptual learning would not entail the epistemological consequences usually expected of theory-ladenness.

A similar question has been asked about cognitive penetration by Lyons (Chapter 3, and Lyons 2011). He argues that in a reliabilist framework, penetrability can sometimes be an epistemologically benign influence because it can enhance the rapid classification of stimuli. While Lyons has in mind synchronic contexts, not the diachronic ones of perceptual learning, his points could apply equally well to the enhancement of
classification by perceptual learning, thus confirming the expectation that perceptual learning should be epistemologically virtuous.

If perceptual learning is epistemologically virtuous despite making perception theory-laden (due to brain plasticity), then some of the parties to the initial cognitive penetrability debate may have been talking at cross-purposes. According to Fodor, encapsulation is an empirically necessary condition for there to be perceptual consensus between observers despite differences among the theories they hold, and failure of encapsulation entails theory-ladenness (see 1.3). But if the dependence of vision on visual memories due to perceptual learning is epistemologically benign, then it may not prevent the perceptual consensus that Fodor sought to secure, despite the fact that it may be a form of theory-ladenness. (Whether the information contained in visual memories can count as a theory of sorts is an issue that we cannot settle here, so we shall leave the question open.)

2.6 Cognitive penetrability and attention

The concept of attention has played an important role since the early days of the cognitive penetrability debate. Fodor (1988) first used the ‘attention-shift argument’ to rebut Churchland’s contention that ambiguous figures, like Jastrow’s rabbit/duck drawing, illustrate the dependence of visual contents on ‘higher cognitive assumptions’ (Churchland 1988: 171). Churchland’s idea was that, depending on which concept one applies, one can see the figure differently each time. Fodor argued that in such cases of visual ambiguity, attention focuses visual resources on different parts of the scene each time; those resources then perform visual processing without any further contribution from attention or higher cognitive assumptions; and consequently, the output of the visual processing differs each time, but not due to any direct influence of concepts on perception. In similar spirit to Fodor, Pylyshyn sought to restrict the influence of cognitive states on perception to ‘the allocation of attention to certain locations or certain properties prior to the operation of early vision’, and the operation of late recognitional procedures which can involve endogenous attention, especially when the scene is ambiguous or sensory exploration is required to understand it (1999: 344; 2003: 90). This classical impenetrabilist view of the role of attention in perception has been amended and developed by Raftopoulos (2009), who distinguished several forms of attention and granted that some of them are part of early visual processing. The classical view is challenged by Mole (Chapter 9). To understand the shape of the current debate, we have first to explain some of the background literature on attention.

Attentional mechanisms are needed because visual scenes typically contain more information than vision can process at any given time; the visual system selects one or a few objects at a time for more thorough processing, as inattentional blindness strikingly illustrates. Attention can bias the competitive interactions among stimuli by increasing the activity of neurons processing some of the stimuli (Desimone and Duncan 1995), for instance by lowering firing thresholds for salient feature detectors or by increasing their activity to enhance their output (Egeth et al. 1984; Kahneman et al.
numerous specific roles have been assigned to attention (see Carrasco 2011 for a review). But two distinctions concerning attention are particularly important for the penetrability debate: the difference between exogenous and endogenous attention, and the difference between the ‘spotlight’ and the ‘biased competition’ views of attention.

Exogenous attention is drawn by the stimulus, or some early processing of the stimulus, and peaks at 100–120 ms; endogenous attention is cognitively driven and takes approximately 300 ms to deploy (Carrasco 2011). Both forms of attention can be overt, covert, spatial, or feature-oriented. (An example of endogenous spatial attention is expecting something to appear at a location and focusing attention on that location; watching out for snake-like shapes on a hike would be feature-based endogenous attention.) Endogenous attention is a cognitive phenomenon on any available criterion: it is personal-level, agentive, and takes so long to deploy that the scene attended to has already been processed by vision beyond the point of object recognition. Exogenous attention, however, can occur in early vision (see e.g. Desimone and Duncan 1995; Vecera 2000; Kravitz and Behrmann 2011). Endogenous effects on early visual processing would amount to CP, while exogenous effects can be compatible with impenetrability of the visual processes in which they are involved.

The ‘spotlight’ (or ‘lens’) conception of attention (Posner 1980; Eriksen and Yeh 1985) construes attention as a distinct faculty or process that intervenes to resolve perceptual competition between representations and select some of them for further processing. The ‘biased competition’ account instead construes attention as the result of a perceptual competition which has been won by some set of stimuli and the way they were processed (Desimone 1998; Duncan 1998; Desimone and Duncan 1995): some properties of the scene just become salient in the course of visual processing. This concept of attention lends itself to use by theories which argue against the direct influence of endogenous attention on early visual processing and locate the influence of attention at an early visual stage potentially isolated from cognitive influence. However, the biased competition account does not have to be used in this way; it does not explicitly preclude unconscious cognitive influences from being among the factors that make certain properties of the scene salient during visual processing. In this sense, the distinction between these two conceptions of attention is potentially orthogonal to the previous one between endogenous and exogenous attention.

While the classical impenetrabilist view of attention, just described, is that attention focuses visual resources and those resources then perform their processing without any further contribution from attention, the opposite view is taken by Christopher Mole (‘Attention and Cognitive Penetration’, Chapter 9). Mole holds that ‘perceptual effects that are owing to what Bruner and Postman call “set or expectancy” can more intuitively be described as effects of attention’. For this, Bruner and Postman’s (1949) perceptual set would have not just to be capable of directing the focus of attention to some set of stimuli already pre-packaged by visual processing, but to contribute to the way in which that processing yields its outputs. This is just what Mole argues,
by using Kravitz and Behrmann’s (2011) study of the role of feature-based attention in vision to show that attention is inextricably involved in early visual processing. In other words, attention is not ‘held off’ while vision does its processing, as the attention-shift argument has it. Mole also enlists Desimone and Duncan's biased-competition theory of attention against Pylyshyn's view of the role of attention in perception. He sees Pylyshyn's view as excluding attention from visual processing, but the competition theory as integrating it into vision: 'The integrated competition theory should not be thought of as a theory that tells us what happens to the perceptual processing of objects as a consequence of attention having been directed at that object. It should, instead, be thought of as a theory of what it is for attention to be directed at an object.' For example, Mole describes findings by Kravitz and Behrmann that semantic information can affect feature attention in early visual processing; on his reading of the biased competition account, this means that semantic information can unconsciously affect the competition for attention by early visual resources.

The biased competition account of attention has been used in a different way by Raftopoulos (2009). He argues that feature attention can be part of early visual processing (not something extraneous to visual processing as Pylyshyn holds), but that in those cases it is not endogenous, nor necessarily unconsciously directed by concepts. In that case, early feature attention could be part of an integrativist cognitive architecture, like that proposed by Burnston and Cohen (Chapter 4), in which perceptual states interact without their interaction entailing cognitive penetration. Raftopoulos made extensive use of the idea that feature attention can be restricted to early visual processing, arguing from Desimone and Duncan's biased-competition account, Rensink's (2000a; 2000b) coherence-field theory of attention, and Vecera's (2000) distinction between object segregation and object individuation, to the conclusion that attention does not directly affect early visual processing. For example, compare the following two scenarios concerning the role of attention in constructing visual objects. If objects were individuated in perception by having attention consciously focused on them, that would mean that object individuation is cognitively driven and cognitively penetrated. But if, when we consciously scan a scene, perception stops at objects which have already been segregated by visual processing in which feature attention plays an unconscious role (Vecera 2000), then object individuation is exogenously driven and could be cognitively impenetrable.

A clear-cut case for the cognitive penetration of vision by attention could be made if we could show that agentively driven attention directly affects processes defined as part of early vision by Pylyshyn. Zeimbekis (Chapter 12) argues that consciously driven spatial attention can affect visual processes that construct volumetric representations from monocular cues. Another attentionally mediated form of cognitive penetration would occur if semantic information determined the focus of attention unconsciously on expected features or locations. For example, Tye (1995: 140) holds that differences in the phenomenal content of perception caused by visually ambiguous figures, such as Jastrow’s duck/rabbit figure, may be unconsciously caused by the conceptual abilities of
viewers. In that case, it could be argued that feature attention is unconsciously focused due to concept possession, making visual content cognitively penetrable through the mediation of attention. However, what Tyeh has in mind is not always easy to distinguish from cognitive influence on late vision. For instance, Raftopoulos (2009: 72) cites evidence of the penetration of object-recognition processes in which, following an unconscious analysis of the semantic content or gist of the scene (in the parietal cortex at about 160 ms), visual resources are directed to the sites in the visual scene where useful information is most likely to be found. Here, conceptually encoded information about the characteristic properties of kinds directs the visual search to enhance and accelerate object recognition, but the effect is on late vision.

Finally, another attention-borne indirect effect of cognition on early vision is pre-cueing. Pre-cueing occurs when a perceiver expects either an unspecified kind of stimulus to appear at a certain location of her visual field, or a specific kind of stimulus to appear anywhere in her visual field. The expectations influence the baseline activations of neurons in the visual areas whose receptive fields fall within the cued location, or of neurons that encode the features of the cued stimulus. These biases are anticipatory and occur before stimulus presentation (Nobre et al. 2102: 161). They are top-down biases because the cue is stored in working memory and activates the relevant neurons in the visual areas through a top-down flow of information. To deny that this amounts to cognitive penetration of early vision, one could try arguing that pre-cueing does not affect visual processing in a direct, online way, but just sets the initial values of certain parameters for subsequent computations. This is the kind of argument that Mole’s contribution, described above, seeks to rebut. (The question of whether indirect cognitive influences on early vision should count as cognitive penetration is discussed in Section 3, ‘Definitions of Cognitive Penetrability’.)

2.7 Perception and cognition: is there a distinction?

For causal relations between two states to count as cognitive penetration of perception, it has to be possible to maintain a distinction between perceptual and cognitive states. The same distinction is presupposed by claims that perception is not, or cannot be, penetrated by cognitive states. As Dretske (Chapter 6) writes, ‘the first step . . . is to carefully distinguish perception from cognition.’ An alternative view would be to see the penetrability question as primitive: the cognitive impenetrability of perception would distinguish perception from cognition, but penetrability would mean that perception cannot be separated from cognition. We think this second approach should be rejected, for the following reasons.

Although statements which describe evidence of cognitive penetration could perhaps be formulated as denials of the perception/cognition distinction, they do not have to be formulated in this way and rarely are today. For example, when describing putative effects of long-term memory on visual colour processing, Hansen et al. (2006: 1368) state that their results show a high-level cognitive effect on low-level
perceptual mechanisms’. To reformulate this as a claim that there is no distinction between the states, Hansen et al. would have had to say that the effects of long-term memory on visual colour processing mean that long-term memory and visual colour processing cannot be distinguished. Instead, they construe the events as effects of one system on another distinct system, keeping the systems apart as relata. A precise proposal about how to preserve distinctions between systems in cases of penetration is proposed by Lyons (Chapter 3): admitting cognitive penetrability would not obliterate the distinctions between information-processing systems because the systems could still be defined on the criteria of isolability, unitariness, and specialization (see 2.1).

It’s worth noting that in the 1940s, evidence about what we today call ‘cognitive penetrability of perception’ was sometimes interpreted in a way that suggested obliteration of the perception-cognition distinction: ‘all perceptual experience is necessarily the end product of a categorization process’ (Bruner and Goodman 1947); ‘the perceptual effect of a stimulus is necessarily dependent upon the set or expectancy of the organism’ (Bruner and Postman 1949, cited by Mole, Chapter 9). Subsequent reactions against such accounts of perception in Marr’s account of vision, and Pylyshyn’s denial of the continuity of cognition and perception—to the extent that it is based on Marr’s hierarchic conception of vision—may be partly responsible for the shift in how penetrability claims are formulated. This takes us to our second point:

Granting the cognitive impenetrability of perception could at best be part of some way of securing a perception/cognition distinction. If an information-processing system is cognitively impenetrable, that implies that it functions autonomously from certain other systems and can be defined without reference to them. But to secure the perception/cognition distinction, some explanation needs to be given of why the first system deserves to be called perceptual and the others cognitive. So we cannot see impenetrability as primitive and make the perception/cognition distinction depend on it; if anything has chances of being primitive in the debate, it is the status of some processes as perceptual and others as cognitive. Briefly, whether one defends CP or CI, one needs to keep perceptual and cognitive states distinct; defending CI additionally requires that certain kinds of relations do not obtain between the states. So we propose to take the perception/cognition distinction as more basic than questions about penetrability and impenetrability.

The perception/cognition distinction is not immune to doubt. An extreme way to doubt it would be to adopt a view of the brain as devoid of any functional hierarchy (see Lashley 1929), but nobody today appears to be tempted by this position. Interactionists, who emphasize the existence of feedback from areas activated later by incoming information to areas which were activated earlier, do not deny that cognition has a functional cognitive hierarchy (Churchland et al. 1994; McClelland and Rumelhart 1986; Vecera and O’Reilly 1998). As Peterson (2005: 189) describes them, ‘these models maintain a hierarchical structure in that lower-level processes must at least be initiated before higher-level processes are initiated.’
More moderately, one might worry, like Shea (2014), that perception is hard to individuate as a category because (i) it depends on top-down influences, Bayesian priors, and cross-modal effects, and (ii) there are borderline states that appear to be neither perceptual nor cognitive.

On the first point, we have seen (Section 2.2) that early vision can be distinguished from cognition on functional grounds by using evidence of dissociations from neurology, and that even late visual processes can be mapped functionally by using such evidence. (For an explanation of how such evidence is used to distinguish items in cognitive architecture, see Mahon and Wu, Chapter 8, §2.) Warrington and Taylor's (1973; 1978) work on left/right parietal lesions was used by Marr (1982) to show that processes which precede viewer-centred brain representations of visual scenes are dissociable from the processes that assist object recognition with semantic feedback. Humphreys and Riddoch's (1987) work on visual agnosia was used by Pylyshyn to show that template matching, which is a late visual process, is dissociable from both early visual computations and semantic information. Mahon and Wu (Chapter 8) construe cognition as a set of semantic information-processing systems which survive damage to specific modalities. Caramazza et al. (1990) argue from clinical evidence for a functional distinction between (a) semantic information which is accessed by object-recognition processes across modalities and even by word recognition and (b) information-processing that is specific to sensory modalities. As Mahon and Wu write, 'In experimental cognitive science, “semantics” is generally operationalized as the information that mediates the mapping from input to output systems, where input and output systems are modality-specific in terms of both their content and their format.' In other words, cognitive science has a working hypothesis about how to distinguish perception from cognition.

There is also a working hypothesis for defining an important part of perception, early vision, by using the timing of neural processes as they spread from sensory surfaces into areas of the brain that function autonomously from areas further downstream. The set of processes that Lamme and Roelfsema call the feedforward sweep and local recurrent processing qualifies as non-cognitive due to the priority of those processes relative to doxastic and personal-level processes, and as visual due to the traceability of the stimulus to the eyes. These seem good enough reasons to call them perceptual.

Shea's second point is that there are borderline states which are hard to classify as perceptual or cognitive. These include Carey's (2009) account of the representation of agency on perceiving animated figures, Michotte's (1963) accounts of perceiving causal relations, and Dienes and Perner's (2003) epistemic feelings. Such borderline cases are the subject of several chapters in this book. Burnston and Cohen argue that systems in charge of representing certain higher-order, multimodal spatial relations (Koechlin et al. 2003) may turn out to be non-modular and isotropic. In that case, it could be argued that such systems count as cognitive. However, on Burnston and Cohen's account, other processes which are standardly thought of as cognitive, such
as recognizing scenes of chasing and recognizing certain social cues (Gao and Scholl 2011; Langton et al. 2000) might turn out to be modular, anisotropic, and thus arguably noncognitive.6

The most prominent borderline cases are perceptual contents. If perceptual experience represents high-level properties such as kinds (Siegel 2010), concepts may either be part of perceptual states or else determine those states’ contents. In that case, can perceptual contents—for instance, conscious visual contents—legitimately be counted as perceptual and not cognitive? One way to respond to this challenge is to try to deny Siegel’s thesis. Dretske (Chapter 6) devises a thought experiment which serves as a criterion to determine what should be excluded from ‘cognitively unspoiled’ perceptual experience. He applies the criterion to Siegel’s claim that kinds are represented in perceptual experience, and concludes that while low-level properties that characterize kinds are represented in perception, kinds themselves are not. (For detailed discussion see Section 4.1. Elsewhere, Dretske has argued at length against the idea that perception and cognition are ‘merely different stages in a more or less continuous information-handling process’; see Dretske 1981: 135.) Another possible response (to the challenge to the perception/cognition distinction from high-level perceptual contents) would be to argue that if perception represents kinds, it does so in some nonconceptual way, and that this prevents conscious perceptual contents from qualifying as cognitive states, construed as thought-like, conceptual states. For example, semantic information could influence vision in a way that made the features which characterize kinds more salient in visual experience, yet without making those contents conceptual (Tye 1995 arguably makes such a proposal). Macpherson (Chapter 13) argues that the cognitive penetration of perception is compatible with the thesis that perception has nonconceptual content; if Macpherson is right, then perhaps perceptual experience can represent kinds and still not be a cognitive state.

Dokic and Martin (Chapter 10) examine another borderline case which sits between perceptual and cognitive states: epistemic feelings during perception. They argue that feelings of familiarity depend on recognitional abilities, and that such feelings can alter the overall phenomenology of perception without affecting perceptual content. They also hold that epistemic feelings of confidence can explain the memory colour effect (Delk and Fillenbaum 1965; Hansen et al. 2006) as occurring at the level of judgment, not sensory experience. If Dokic and Martin are right, epistemic feelings experienced during perception are cognitive states, not perceptual ones.

We do not think that the borderline cases threaten the perception/cognition distinction, because they mainly affect the classification of late visual processes as perceptual or cognitive. Since there are well-established working hypotheses about how to keep early vision distinct from cognition, it seems more reasonable to expect that perception

---

6 For more on cross-modal effects, see Section 2.1. We have already explained why another potential set of borderline cases—hard-wired formation principles and Bayesian priors—do not count as cognitive processes or states (Section 2.3).
and cognition can be distinguished if only on the grounds provided by those hypotheses; and to continue to make statements about the cognitive penetrability or impenetrability of perception in the expectation that the distinction will be upheld.

But if the borderline cases do not refute the perception/cognition distinction, they are definitely an issue for the penetrability debate. Suppose that definitions of early vision in functional terms or in terms of processing latencies succeed in isolating a set of non-borderline perceptual processes; then key issues for the penetrability debate will be played out from the stage that Marr called 'late vision' onwards, where the perception/cognition distinction is much harder to trace. While functional dissociations do allow us to isolate some late visual processes (like template matching in recognition—Pylyshyn 2003: 72), late vision is much harder to distinguish from cognitive states like propositional attitudes, concepts, long-term memories, imagery, or cognitively driven attention. Processes at these stages are involved in object recognition, perceptual belief, demonstrative thoughts and their verification, long-term visual recall, and visual imagery. They are important for epistemological reasons, and their cognitive penetrability would still affect perception's experiential outputs, to which subjects have access for reports and judgments.

3 Definitions of Cognitive Penetrability

A necessary condition for the cognitive penetrability of information-processing systems was put forward by Pylyshyn by using the concept of semantic coherence: ‘if a system is cognitively penetrable then the function it computes is sensitive, in a semantically coherent way, to the organism’s goals and beliefs, i.e., it can be altered in a way that bears some logical relation to what the person knows’ (1999: 343). Applied to perception, semantic coherence would be a relation between the contents of perceptual and cognitive systems. Macpherson (Chapter 13) provides a clarification of the nature of such a relation: while the content of a visual state may be nonconceptual and the content of a cognitive state conceptual, a semantic relation between contents can hold if concepts that specify the content of the conceptual state are required for a correct specification of the content of the visual state.7

In addition to the content-based condition, Pylyshyn placed a requirement on vehicle states: for there to be CP, the perceptual system has to draw directly on the informational resources of a cognitive system in performing its computations. This condition is implicit in Pylyshyn’s denial that attentionally mediated cognitive influence on perception qualify as CP (1999: 344; 2003: 90). It can be used to disqualify several key case types from counting as CP even though they meet the content condition:

7 Stokes (Ch. 2) considers an attenuated interpretation of Pylyshyn’s semantic criterion, which loosens it to ‘representational coherence’: ‘This is the essence of what we mean by cognitive penetration: it is an influence that is coherent . . . when the meaning of the representation is taken into account’ (Pylyshyn 1999: 365, cited by Stokes).
(a) Concept possession can determine the focus of spatial or feature attention, with the result that which stimuli are processed visually depends causally and counterfactually on a cognitive state in a way that is semantically coherent with the content of that state. Such selective allocation of attention can occur either following stimulus onset or prior to stimulus onset as a result of pre-cueing (see Section 2.6). For example, conceptually encoded information about the shapes or familiar orientations of kinds can affect figure–ground segregation, or which visual experience we have of an ambiguous figure. Pylyshyn denies that this dependence amounts to cognitive penetration on the grounds that visual processing is only affected indirectly by cognition, through attentional selection of the stimuli to be processed, not by drawing directly on the cognitive state’s computational resources. (b) As explained in Sections 2.4 and 2.5, local top-down influences on incoming visual information can be indirect bearers of theoretical commitments through perceptual learning. Here again, the semantic coherence criterion is met but the cognitive influence is mediated by an external historical causal link between cognitive and perceptual vehicle states. The link is more attenuated than in Pylyshyn’s attentionally mediated cases, so on Pylyshyn’s account it should not count as a form of CP. Commentators who are prepared to call cognitive influence through perceptual learning a form of CP call it ‘diachronic penetration’ (Churchland 1988; Stokes, Chapter 2). (c) The indirect causal link between cognitive and perceptual vehicle states could also be internal and still not secure cognitive penetration on the vehicle criterion, as suggested by an example from Macpherson (Chapter 13):8

Suppose that Murdo believes that aliens are attacking Earth. This belief causes stress, which induces a migraine. Suppose that whenever Murdo has a migraine, he experiences flashing lights in the top half of his visual field. So suppose that Murdo visually experiences flashing lights in the sky on account of having his migraine. The content of this experience bears a semantic relation to the belief that caused it—the belief about the alien attack—but one might want to deny that this is a case of cognitive penetration.

Wu (2013) similarly places both semantic and vehicle-based requirements on CP. According to Wu, failure of informational encapsulation would be entailed by a conjunction of the following three conditions:

1. Internal causal link: S’s visual experience V with content p is causally dependent on a nonvisual system Y via an internal causal link.
2. Computational condition: The influence of Y on V(p) makes the visual content intelligible owing to the computations that underwrite p using information in Y as a resource.
3. No explanatory defeaters: The resulting p is not explained by changes (i) in the proximal stimulus, (ii) the state of the eyes, (iii) the locus of attention.

8 This case, ‘Murdo’s migraine’, should be distinguished from the migraine described in Macpherson (2012), which does not preserve semantic coherence.
The computational condition is semantic as well as being about vehicle states: the computations that support the visual state have to be able to exploit the conceptual representational resource. (This entails that late vision is cognitively penetrable since, for example, object recognition processes use conceptual information.) Jointly, (1) and (2) exclude perceptual learning from counting as cognitive penetration, since in perceptual learning the causal link is external. With (3.iii), they exclude attentional effects from being cases of cognitive penetration—but with the exception of pre-cueing: pre-cueing causally affects perceptual processing through an internal mental link, and would count as cognitive penetration on Wu’s account.

Vehicle-based and semantic conditions are also required by Macpherson (2012). Macpherson describes cognitive penetrability as a situation in which ‘that which is perceived, the viewing conditions, and the state of the sensory organ are held fixed; the location of one’s attentional focus is held fixed; different perceptual experiences are caused on account of differences between the states of cognitive systems; and the difference between cognitive states makes this difference intelligible’ (Macpherson 2012: 28, 29). Intelligibility satisfies Pylyshyn’s content criterion, and the condition that fixes attentional focus precludes at least external indirect causal connections. Nevertheless, Macpherson finds an important case type that satisfies the definition: the influence of memory colour and typical colours represented by concepts (Delk and Fillenbaum 1965; Hansen et al. 2006; Perky 2010; Segal and Fusella 1991).

Other commentators weaken the semantic condition: they do not require any relation of semantic intelligibility to hold between the contents of penetrating and penetrated states as long as a cognitive state causes changes to perceptual content. For instance, Siegel (2012) defines cognitive penetrability as the nomological possibility that cognitive or affective states can cause a change in the visual contents that are or would be experienced while seeing and attending to the same distal stimuli under the same external conditions. According to Stokes (2012), a perceptual experience E is cognitively penetrated if and only if (1) E is causally dependent upon some cognitive state C and (2) the causal link between E and C is internal and mental. Both definitions include a requirement that the content of penetrated perceptual states be different to what it would have been without penetration, and both include a vehicle-based requirement: Siegel’s through attention-fixing, Stokes’s by excluding external causal links. But neither requires that the content of penetrating states make the content of penetrated states intelligible, or that they be semantically coherent.

One motivation for dropping the semantic intelligibility condition is that some apparently paradigmatic cases of cognitive penetration do not seem to satisfy it (see Stokes, Chapter 2). Consider wishful seeing (Balcetis and Dunning 2010), in which subjects see a gift card with a $25 balance as being closer to them than one with the $0 balance. A conceivable explanation of wishful seeing is that, as Lyons puts it (Chapter 3), ‘thinking they’re going to get money makes subjects happy, and this
causes biochemical changes that affect perception in systematic ways, making objects look closer, hills less steep. It is not clear how such a relation between penetrating and penetrated states could be described as semantically coherent. Another reason to drop the condition is that penetrating states may not always be conceptual states. If and when they are not, it may not be possible to define a relation of semantic intelligibility between penetrating and penetrated states. Consider the following cases:

(1) If knowing what I am eating directly affects its taste, that is a case of concepts penetrating perception; but if knowing what I am eating causes feelings (such as disgust—see Nanay 2012) which affect taste, then the penetrating state is an affective one which may not be conceptual.

(2) Cognitively driven spatial attention (see 2.6 on attention) is a potential source of cognitive penetration that does not involve concepts: while it is frequently driven by conceptual information (e.g. when we focus attention spatially as a function of conceptual expectations about the orientation of objects), perhaps endogenous spatial attention can also be driven by habits of sensory exploration or by some form of perceptual activity that involve no specific concepts.

However, it is worth noting that the semantic intelligibility condition does apply to many key cases of penetrability currently under discussion. Cases on which much of Siegel’s work has focused, in which possessing a recognitional or a sortal concept affects which visual experience one has, would preserve semantic coherence. Whether those effects are unconscious effects on early vision or conscious ones during late vision, they alter visual phenomenology in ways that depend on semantic information about the characteristic features of kinds. An intelligibility condition is central to Macpherson’s (Chapter 13) account of why cognitively penetrated perceptual states can still have nonconceptual contents: the nonconceptual contents of penetrated states are specifiable in terms of concepts that specify the contents of the penetrating states. Finally, Macpherson’s (2012) thesis about the penetrability of colour perception also preserves semantic coherence: it claims that penetrated colour experiences represent actual colours as being closer to typical colours—and information about typical colour is semantic information.

Another recent tendency is to doubt the need for a direct relation between vehicle states. For example, Macpherson (Chapter 13) makes the vehicle condition optional, thus potentially admitting as cases of CP all the attention-shift cases that preserve the intelligibility condition (enumerated earlier in this section: pre-cueing, ambiguous figures, perceptual learning, and internal contrived cases like Murdo’s migraine). Stokes (2012) sets a condition on vehicle states that excludes attention-shifts from counting as cases of cognitive penetrability by stipulating that the penetration relation must be internal and ‘involve mental states and processes’; however, he attenuates the condition by adding ‘but with no restriction on how long that causal chain is’. Stokes (2012) also drops the intelligibility requirement. This qualifies as cognitive penetration the wishful-seeing cases, certain indirect forms of orectic and affective influence, and
possibly pre-cueing (which could be counted as internal because it is a lowering of the firing thresholds of neurons to make them more receptive to certain stimuli).

What fuels these doubts about the vehicle condition, and the ease with which it disqualifies attention-shift cases, is a recent consequentialist conception of cognitive penetrability (the expression is used by Stokes, Chapter 2). Deniers of the cognitive penetrability of perception constantly rely on the attention-shift argument, but their opponents increasingly point out that this tactic has an epistemological cost: applying the vehicle condition succeeds in defining parts of visual processing that are isolated from direct cognitive influence, but indirect cognitive influence can be just as harmful epistemologically. It can also defeat one of the main objectives of defending the hypothesis in the first place.

For example, Lyons (Chapter 3) compares two potential explanations of wishful seeing. On one explanation, wishful seeing is an attentional effect; on the other, it is 'the direct cognitive penetration of vision by desire: subjects want the item to be closer, and this makes it look closer.' Direct penetration minus the attentional mediation would satisfy Pylyshyn's vehicle condition and ensure transfer of information between vehicles. But Lyons concludes that whether there is direct penetration or indirect attention-mediated influence, the outcome is the same epistemologically, especially for the perceptual justification of belief. This suggests that tightening the definition of penetrability to exclude attentionally mediated cognitive influences on perception may technically secure impenetrability, but still admit the epistemological consequences that the impenetrability hypothesis was meant to avoid. (For more on this topic see Section 4.7.)

Dustin Stokes (Chapter 2, 'Towards a Consequentialist Understanding of Cognitive Penetration') takes the disputes over the content and vehicle conditions as his starting point, and proposes to settle them by offering a consequentialist definition of cognitive penetrability. He argues that there is no consensus today on what cognitive penetrability is, with the result that the penetrability and impenetrability hypotheses have in effect become empirically untestable. Stokes shows that two standard criteria for penetrability yield contradictory and ambiguous verdicts in cases whose interpretation should be decisive for the debate. The first is Pylyshyn's semantic coherence; the second is Stokes's own (2012) criterion of internal causal influence. The definitions yield conflicting verdicts on key case types: the influence on size perception of beliefs about the value of perceived objects, and diachronic influence on perception which enables perceptual representation of kinds according to Siegel. The first cases do not amount to cognitive penetration on Pylyshyn's criterion, while they do on Stokes's (2012) earlier criterion. On the other hand, the perception of kinds does amount to cognitive penetration on Pylyshyn's criterion, while other cases of diachronic influence on perception may not count as penetration on the weaker criterion of internal causal influence. Stokes then puts forward a consequentialist definition of the cognitive penetrability of perception based on the existing consensus about its presumed consequences: theory-ladenness of empirical observation, vitiation of the
epistemic role of perception, and denial that perception is modular. On the consequentalist definition, cognitive influence on perception through attention-shifting qualifies as cognitive penetration for epistemological reasons, and perceptual learning amounts to cognitive penetration because it lends a degree of theory-ladenness to observation.\(^9\)

To summarize, two conditions were defined for the cognitive penetrability of perception in the original computationalist framework: (1) the perceptual system draws on the informational resources of a cognitive system to perform its computations (the vehicle condition); (2) it generates output representations that it would not generate if it did not draw on the cognitive resources, and those representations stand in a relation of semantic coherence to the penetrating informational resources (the content condition). Gradually, proposals have emerged to weaken or drop these conditions. In particular, effects of cognitive influence on perceptual content, but without semantic coherence, are now widely seen to qualify as cases of penetrability. This is despite the fact that apparently good cases of cognitive penetrability can be found even by keeping the coherence condition.

4 The Philosophical Relevance of Cognitive Penetrability

The concepts outlined in the previous sections are constitutive of the topic of cognitive penetrability. In this section, we explain how hypotheses about penetrability and impenetrability affect the outcome of philosophical debates in several areas: the contents of perception, perceptual and cognitive phenomenology, nonconceptual content, consciousness and awareness, representationalism and realism, perceptual warrant, and action.

4.1 High-level and low-level perceptual contents

Debates about the contents of perception oppose the view that perceptual experience represents low-level properties (Dretske 1981; Tye 1995; 2000; Clark 2000; Raftopoulos 2009) to the view that it represents high-level properties and especially kinds (Siegel 2006; 2010; Bayne 2009; Masrour 2011). Part of this debate uses the methodology of phenomenal contrast cases, in particular for patients with associative agnosia, who cannot classify objects under sortals even though their ability to perceive low-level properties remains intact. Bayne (2009) has argued that such patients, unlike normal subjects, suffer from an inability to perceive kinds. Use of the contrast methodology

---

\(^9\) A contrasting view is held by Raftopoulos (2001a; 2001b; 2006; 2009; forthcoming), who holds that indirect effects on perception are not epistemologically damaging in the way that direct effects are. In that case, if the term ‘cognitive penetrability of perception’ is reserved for cognitive influences that admit consequences such as theory-ladenness, Kuhnian incommensurability or constructivism, only direct cognitive effects on perception should count as cases of cognitive penetrability.
(Kriegel 2007; Siegel 2010) in cases like the auditory perception of sentences (Strawson 1994), the perception of kinds, or cases of agnosia also suggests that perceptual contents can be partly a function of semantic information about the objects of perception. (A synopsis of claims about high-level content and their methodology is provided by Briscoe, Chapter 7, §3.)

High-level perceptual content would imply cognitive penetration in some form, whether of early vision or late vision. Thus, one way to defend low-level content is to argue that perception is not penetrable in the ways that would be required for subjects to have high-level perceptual contents. This would involve showing that mechanisms which generate perceptual representations of low-level properties (orientation, movement, texture, shape, brightness, pitch, and so forth) function independently of semantic information about kinds; a position defended by Marr (1982) as well as by proponents of the impenetrability of early vision. Additionally, the content generated by encapsulated processes would have to be a level of content that subjects can be aware of, since otherwise it cannot be the content low-level theorists have in mind. Defenders of high-level content, on the other hand, want to know whether their theses imply that perception is cognitively penetrated and whether that would make their account of perceptual contents incompatible with certain accounts of perceptual warrant (see the contributions to this volume by Lyons and Siegel (Chapter 3 and Afterword; Siegel 2010, 2012; Lyons 2011).

Reciprocally, the question of which properties are represented in perception affects the penetrability debate. Evidence from agnosia for high-level perceptual content, or support for it from the contrast methodology, obliges defenders of impenetrability to show that high-level content is caused by cognitive penetration only of late vision. This in turn raises a question for both the impenetrability hypothesis and the low-level view of perceptual content: how to account for later visual processes like object recognition. On the impenetrability hypothesis, late visual states are affected by concepts; but being visual, they must also qualify as perceptual states. So how can the contents of perception be low-level properties?

One way to deny that perception represents kinds and still account for Siegel's (2010) contrast cases may be to argue that concept possession causally affects early perceptual processing, but in a way that keeps conscious perceptual content low-level. For instance, it could be argued (see Tye 2000: 61) that when we speak a language, the contents of our auditory perception of sentences include salient word boundaries which would not be represented if we did not speak the language, but without including recognition of the words.

Some accounts of low-level content are closely connected to a defence of the cognitive impenetrability hypothesis. Dretske's (1981) defence of low-level content is part of his denial that perception and cognition are 'merely different stages in a more or less continuous information-handling process', that they are 'mutually interdependent and cannot be separated except by arbitrary rules', and that 'recognition, identification, and classification... occur at every phase of the perceptual
process’ (1981: 135, 254). If the contents of perception are counterfactually dependent on causal cognitive influences, then it seems that low-level contents result from classification and that there is continuity between perception and cognition after all.

In his contribution to this volume, Fred Dretske (Chapter 6, ‘Perception versus Conception: The Goldilocks Test’) argues against high-level perceptual content by devising a thought experiment, the ‘Goldilocks test’. This thought experiment is deceptively simple: the more one thinks about it, the more robust it turns out to be. In the thought experiment, subjects endowed with idealized painting skills can share their perceptual experiences by painting them. Thus, subjects can have each others’ visual perceptions and even superimpose them for comparison. This imaginary scenario is then applied to Siegel’s (2006; 2010) main contrast case, in which Siegel claims that acquiring the recognitional concept for pine trees alters the visual experience caused by pine trees. Call the subject who possesses the recognitional concept pine tree ‘E’ (for expert), and the subject who does not possess the concept ‘N’ (for novice). Now, imagine that the subjects paint a pine tree according to their visual experiences of it, and thus display and superimpose their visual experiences for comparison. If concept possession alters visual experience, N’s picture should be different to E’s picture. There are two ways this could happen:

1. N’s picture does not cause the visual experience of pine-tree-ness because the picture contains too little information. In that case, E should be able to add the information to N’s picture. But this is excluded by the thought experiment’s premises, which state that E and N have the same stimuli and saccades when they look at the pine tree. One may ask if subjects, saccades and attention do operate identically when only one of them possesses a recognitional concept for the object perceived. But the thought experiment legitimately brackets this difference on the grounds that the stimuli are available to be saccaded or attended by subjects who possess or do not possess the concept alike. To deny this premise of the thought experiment, the defender of high-level content would have to commit herself to the unlikely view that if you don’t possess the concept, then you cannot visually experience certain shapes and colours.

2. N’s picture does not cause the visual experience of pine-tree-ness because the picture contains too much information. In that case, the picture would represent the features of a kind more specific than being-a-pine-tree, such as being-a-White-Pine specifically. But, as Dretske points out, ‘If one is going to see kind properties, one should be able to see them despite seeing details distinctive of more specific kinds’, and this makes it hard to explain how N’s picture could fail to cause an experience of pine-tree-ness in E by having too much information. Dretske concludes that N’s picture will have no more, and no less, information than E’s picture, but the same information. Since the pictures of the thought experiment are the subjects’ visual experiences of the object, their visual experiences will also have the same information. What does differ, according to Dretske, is that N just doesn’t know that arrangement of coloured shapes is a realization, a perceptually
Robert Briscoe (Chapter 7, ‘Cognitive Penetration and the Reach of Phenomenal Content’) analyzes the three methods which have been used to reach the rich content view: phenomenal contrast (Siegel 2010); application of the contrast method to patients suffering from associative agnosia (Bayne 2009); and empirical evidence from attention, detection times, and visual adaptation (used by Fish 2013). He concludes that none of the three methodologies has succeeded in showing that visual experience can come to represent high-level properties as a result of influence from cognitive systems involved in object recognition. For instance, while visual searches are influenced cognitively through saliency maps (Wolfe et al. 2004), this affects the appearance of low-level properties without implying that high-level properties become part of phenomenal contents. Briscoe points out that Rubens and Benson’s (1971) accounts of functional deficits in associative agnosia do not imply changes to the phenomenal character of experience (as claimed by Bayne 2009). (Remember (2.2) that work on visual agnosia has also been used to deny cognitive influence on vision; see Humphreys and Riddoch 1987: 104 and Pylyshyn 2003: 72.) When Fish appeals to evidence about detection times, he mistakenly assumes that the detection of high-level properties means that the properties have been experienced in the phenomenal character of perception even when the exposure time is extraordinarily short (20 ms). Briscoe concludes that, for now, there is no evidence that cognitive systems involved in object recognition penetrate vision causing high-level properties to be represented in phenomenal contents.

The visual experience of low-level properties can be affected by information originating from elsewhere outside the visual system, according to Briscoe. He describes psychophysical and neuroscientific evidence for what he calls ‘synchronic low-level informational penetration’ (SLIP): the effect of haptic information about slant on the visual appearance of surfaces (Ernst et al. 2000), the ‘sound-induced flash illusion’ (Shams et al. 2000), and the proprioceptive scaling of size (originally described by Malebranche). Finally, Briscoe describes what appears to be a genuine case of cognitive penetration that affects visual experience. He argues that findings by Vishton et al. 2007 on the Ebbinghaus illusion, jointly with the immediate way in which the dorsal visual stream is thought to operate, imply that conscious distal intentions penetrate visual size perception. (This claim is discussed again in 4.6.)

4.2 Cognitive phenomenology, perceptual phenomenology, and epistemic feelings

The distinction between cognitive and sensory phenomenology can be used to defend cognitive impenetrability by restricting the phenomenology of representing high-level properties to cognitive, non-sensory processes. According to Marr (1982), object recognition is subject to top-down cognitive influences, while the visual processes...
that lead up to the 2½D sketch are not (see 2.2). Suppose, for the sake of argument, that Marr is right. Then, the absence of top-down influences on the earlier visual processes would allow us to claim that their phenomenology does not represent kinds. But that would not preclude there being a distinctive way it feels to represent kinds. This distinct phenomenology could supervene on later visual processes; it would be part of what Lyons (2011) describes as 'late experiential effects' which do not affect the outputs of sensory processes. This would constitute a case of what Kriegel (2007: 118) calls 'phenomenological inflationism' (the tendency to claim that higher-level properties as such are phenomenologically manifest): the mental representation of kinds, associated with recognitional concepts, would have its own phenomenology in addition to the phenomenology of perceiving low-level properties. In that case, the inability of patients suffering from pure associative agnosia (discussed by Bayne 2009) to represent kinds would be due to a deficit of the recognitional processes that subserve a cognitive phenomenology; this would be consistent with the fact that the visual processes which generate representations of low-level properties remain intact in those patients.

Another way to seek to deny that experiences affected by cognition are sensory experiences would be to show that they are non-sensory epistemic feelings. Jerome Dokic and Jean-Remy Martin (Chapter 10, ‘Looks the Same but Feels Different: A Metacognitive Approach to Cognitive Penetrability’) use recent findings about metaperceptual feelings (e.g. feelings of presence or reality, familiarity, and confidence) to argue that while the possession of recognitional concepts does affect perceptual phenomenology globally, it does not influence visual contents, and that therefore the contrast cases do not imply the cognitive penetrability of vision. Metaperceptual feelings are presented through pathological cases (derealization disorder, Capgras syndrome, and Fregoli delusion) in which the feelings are either missing or over-generated. (For reviews of the literature on metacognition see Dokic and Martin, Chapter 10; Koriat 2007; Dokic 2012.) Dokic and Martin argue that the experience of kinds can be explained by metacognitive epistemic feelings correlated with concept possession, but generated without the concepts entering the contents of perception, and without tokening of the concepts in beliefs. If their hypothesis is correct, then the part of perceptual experience that changes as a function of concept possession is not a representational part that affects the contents of perception, and should instead be factored into psychological modes adopted towards perceptual contents.

Siegel (2010) has considered the possibility that the experience of high-level properties in perception is due to epistemic feelings of familiarity, whether non-representational ones which do not alter visual contents, or those that accompany commitment-involving doxastic attitudes. She objects to the first alternative on the grounds that feelings of familiarity emerge in connection with some content; and to the second on the grounds that the experience of kinds emerges independently of whether a commitment-involving attitude is adopted toward the object of perception. Dokic and Martin, citing recent work on metacognition (especially the work of Koriat
2000; 2007), argue that the feeling of familiarity is ‘grounded on the subpersonal monitoring of the quality of perceptual processes’. In that case, while occurrence of the feelings would presuppose possession of the recognitional concept (or at least repeated perceptions of members of the same kind), it would not require the relevant kind to be represented in perceptual contents. (Koriat 2007 similarly describes as ‘contentless’ the mnemonic cues that affect the fluency of information retrieval and encoding.) Dokic and Martin also apply findings about metacognitive feelings to claims that colour perception is affected by memory colour (Delk and Fillenbaum 1965; Hansen et al. 2006; Macpherson 2012). In these experiments, subjects are placed in situations of what Tversky and Kahneman (1974) call judgmental uncertainty (see Zeimbekis 2013). Dokic and Martin argue that epistemic feelings of confidence play a decisive role in the outcome of matching or colour categorization tasks. Consequently, whether perceived colour is cognitively penetrated will depend on whether such feelings can influence judgment while leaving experience intact.

To show that any experiences affected by cognitive states are not sensory but cognitive, one also has to deal with the existence of cognitive states whose phenomenology is hard to distinguish from sensory phenomenology, such as visualizing and mental imagery. Both sides of the imagery debate consider visualizing and mental imagery to be cognitive states (Tye 1991; Pylyshyn 2007, 2003; Kosslyn 1994; see also Marr 1982), yet the phenomenology of those states has distinctly visual ingredients. The effect of conscious acts of visualizing on the visual experience of perceived shape is studied by John Zeimbekis (Chapter 12, ‘Seeing, Visualizing, and Believing: Pictures and Cognitive Penetration’). Zeimbekis describes two forms of visual ambiguity that involve visualizing. Visually ambiguous figures, especially pictures, have always been used to test the impenetrability hypothesis (Fodor 1988; Churchland 1988; Pylyshyn 1999; 2003; Macpherson 2006; Raptopoulos 2009; 2011). But two forms of ambiguity remain unexplored in the penetrability literature and pose a significant threat to the impenetrability hypothesis: the ability to visually experience a picture surface as flat after it has caused nonconceptual contents representing volumetric shapes; and the ability to use a surface initially perceived as flat to visualize three-dimensional scenes. In both cases, the visual processes which extract viewer-centred volumetric shapes (equivalent to Marr’s 2½D sketch) have to rely solely on monocular depth cues in the absence of parallax and stereopsis. Zeimbekis argues that those processes can be cognitively penetrated by acts of visualizing, including ones that draw on conceptual information about kinds. However, he holds that the penetrability of the visual processes does not weaken their ability to provide perceptual warrant and justification. When pictures support 2½D visual experiences through cognitively driven acts of visualizing, the penetrated visual states do not cause beliefs because they are consciously sustained acts which have a different phenomenology to visual perception. On the other hand, naturalistic pictures with good monocular depth cues, which generate 2½D contents without any need for visualizing, cause visual experiences which are harder to distinguish from object perceptions. Zeimbekis argues that such
states differ from visual object perceptions because they represent volumetric shape without either stereopsis or parallax, and do not engage the motion-guiding dorsal visual stream (see Matthen 2005). A comparison with monocular and binocular object perception suggests that such picture perceptions also differ phenomenologically from object perceptions (for reasons pointed out by Sacks 2010 and Briscoe 2008). Thus, while the visual ambiguity caused by penetration of the visual processes that assign shape can lead to states with contradictory visual contents, at most one of those states can ground a perceptual belief and be taken into account in perceptual justification.

4.3 Nonconceptual content

The hypotheses that early perceptual processes are cognitively impenetrable or informationally encapsulated were formulated independently of the hypothesis that perceptual content is nonconceptual. (See Evans 1982; Peacocke 1992; a similar thesis is developed in Dretske 1969; 1981. For an overview of nonconceptual content, see Bermúdez and Cahen 2011 and Macpherson, Chapter 13.) Yet the concepts of cognitive impenetrability and nonconceptual content appear to be closely related. At first sight, it appears that the cognitive impenetrability of perception could suffice for perception to have nonconceptual content: if perceptual representational states are generated by cognitively impenetrable perceptual processes, then they should not be conceptual states. Conversely, one might ask whether the cognitive penetration of perception would make perceptual states conceptual.

Both questions are important because if impenetrability was a necessary condition for perception to have nonconceptual content, and if the penetrability of perception was established, then the contents of perception would not be nonconceptual. Macpherson (Chapter 13), who argues that perception is cognitively penetrable (Macpherson 2012) but also defends nonconceptualism about perception (2006, and Chapter 13), seeks to show that the two theses are compatible. On the other hand, Raftopoulos (2009; Raftopoulos and Muller 2006) has argued that cognitive impenetrability is necessary for perception to have nonconceptual content.

Raftopoulos (2009) reasons as follows. On Pylyshyn’s hypothesis, memory and conceptual information can influence recognitional processes, and visual processing is impenetrable up to and including viewer-centred object representations. Recent evidence from visual neuroscience and studies of awareness and attention (reviewed in Raftopoulos 2009) suggests that the cognitively impenetrable part of vision does not go beyond the viewer-centred representations which are thought to constitute the content of the representational brain states of which we can be phenomenally aware. The upshot is that, on current defences of the impenetrability hypothesis, the impenetrable part of vision is that whose outputs are representations corresponding to descriptions of perceptual nonconceptual content (viewer-centred volumetric representations, according to Evans 1982 and Peacocke 1992). Raftopoulos construes nonconceptual contents to include both subpersonal computational states and personal-level states.
(like Bermúdez 1995). Since impenetrability would be the absence of transmission of conceptual information to the perceptual processes that yield phenomenal content, it would guarantee the nonconceptual character of a certain part of perception. This is especially important for nonconceptual content conceived as a solution to the grounding problem and as part of the explanation of concept acquisition, since what does the grounding and explaining of conceptual states in these cases cannot be a conceptual state. It also has repercussions for the epistemological questions related to penetrability (see Sections 4.5 and 4.7).

Whether impenetrability is a necessary condition for nonconceptual content depends on two sets of factors. The first has to do with the nature of cognitive influences on perception. It has been argued that some perceptual contents are generated partly under cognitive influence but remain nonconceptual (e.g. figure–ground segregations; see Tye 1995; Peterson 2005; Zeimbekis, Chapter 12). Such cases have been used to argue that ‘it does not seem to be a necessary condition on the nonconceptual content of representations in general that they be insulated from the propositional attitudes’ (Bermúdez and Cahen 2011).

The second set of factors are the definitions one adopts of mental content in general, of perceptual content, and of nonconceptual content in particular. Fiona Macpherson (Chapter 13, ‘Cognitive Penetration and Nonconceptual Content’) sets out to determine whether impenetrability is a necessary condition for perception to have nonconceptual content, by taking into account concept-based and non-concept-based accounts of content, four definitions of nonconceptual content, and two definitions of cognitive penetrability. One of the aims of Macpherson’s chapter is to show that the mechanism of cognitive penetration described in Macpherson (2012) allows us to preserve the thesis that perception has nonconceptual content. In what Macpherson calls ‘classic cognitive penetration’ of perception, the subject has a perceptual state with a content that she could not have if her perception was not causally affected by some cognitive state. In what Macpherson calls ‘cognitive penetration lite’, the subject has a perceptual state with a content that she could have even if her perception was not causally affected by some cognitive state. Macpherson argues that the second form of penetrability is compatible with standard definitions of nonconceptual content. For example, according to the state view of nonconceptual content (Heck 2000), a mental state is nonconceptual when it is possible for the subject to have that state even if she does not possess the concepts that would be involved in a correct specification of the state’s contents. Thus, if one of the content-specifying concepts did causally influence perceptual mechanisms to yield the same content, it would still remain true of the content that we could have had it without possessing the concept, satisfying the state definition of nonconceptual content. But in that case, Macpherson’s definition of ‘cognitive penetration lite’ would also be satisfied.

According to another view of nonconceptual content, the content view, mental contents are nonconceptual to the extent that they do not count concepts as constituents. As Macpherson points out, one way to meet this condition for penetrated perceptual
contents is to adopt a possible-worlds account of mental contents, on which no mental content has concepts as constituents. In fact, Macpherson’s (2012) account of the mechanism of penetration suggests that in a range of cases, the contents of penetrated states are of the same kind as the contents of non-penetrated states; this applies to cases in which visual imagery and perception contribute jointly and indistinguishably to the phenomenology of an experience (Perky 1910; Segal 1971; Segal and Fusella 1971; Chand and Murthy 2007).

Macpherson’s ‘cognitive penetration lite’ is incompatible with one interpretation of Bermúdez and Cahen’s (2011) account of nonconceptual content, and with Raftopoulos and Muller’s (2006) account of nonconceptual content, according to which a content is nonconceptual only if it is not cognitively penetrated. On the other hand, what Macpherson calls ‘classic cognitive penetration’ (in which the subject has a perceptual state with a content that she could not have if her perception was not causally affected by some cognitive state) is incompatible with all four accounts of nonconceptual content. Thus, it turns out that whether the contents of perception are nonconceptual depends on whether all forms of cognitive penetration function in the way Macpherson’s ‘classic cognitive penetration’ does.

4.4 Awareness and attention

The core questions concerning cognitive penetrability and attention were explained earlier (Section 2.6): the timing of different kinds of attention has to be understood before we can know which stages of visual processing they affect and whether they amount to cognitive penetration of perception. Beyond these questions, attention has been assigned a number of philosophically important tasks: accounting for consciousness (Campbell 2002; Lamme 2003; Kriegel 2005; Block 2007b; Mole 2011), explaining how we acquire coherent object representations (Treisman 1993; Vecera 2000), verifying demonstrative thoughts, securing conscious awareness of objects (Dehaene et al. 2006; Lamme 2003), and others. For instance, if the individuation of objects was determined through mediation of feature attention or spatial attention driven by sortal concepts, that would support the form of sortalism Campbell calls the ‘delineation thesis’. Distinguishing different kinds of awareness and attention, the latencies at which they occur, and the processes and states which are thought to be their neural correlates is also crucial for an assessment of hypotheses about the contents of perception and cognitive phenomenology.

The question of whether awareness as a form of consciousness presupposes attention is inextricably tied to the question of whether early vision is affected by attention. On one view, local recurrent processing (see 2.2) gives rise to phenomenal awareness, marking off unconscious brain representations from representations of which we can have some form of consciousness (Lamme 2003; 2005; Block 2007a; 2007b; Super et al. 2003). For this to be consistent with denials that early vision is attentionally modulated, awareness should not presuppose attention. Traditionally it was thought that awareness does presuppose attention (for recent defences see Dehaene et al.
If that is so, then attention must intervene in early vision, potentially modulating it. On the other hand, there is now also empirical evidence to suggest the contrary (van Boxtel et al. 2010; Koch and Tsuchiya 2006; Lamme 2000; 2003; 2005).

These questions about awareness and attention directly affect the outcome of the debate outlined in Section 4.1 about the contents of perception, since those contents are typically thought to be divulged phenomenally. Pylyshyn (2003; 1999) is sceptical about how well introspection can inform us of the contents of different levels of perception, claiming that we introspectively confuse the panoramic conception of the visual scene which is the result of late vision with earlier stages. Raftopoulos, applying Lamme’s distinction between phenomenal and report awareness, holds that ‘the phenomenal content of perception differs from the content of the same experiences when consciously accessed’ (2009: 163). The contrast methodology, used to settle disputes based on introspective reports, may not be in a position to overcome the host of objections raised by Pylyshyn and may not curbs finely as accounts based on the neural correlates of different levels of awareness. For instance, a case has been made that phenomenal awareness, as understood by Block (1996), is the output of local recurrent processing (described in 2.2). Unless neuroscientific findings show that such processing can interact with long-term memory, the differences in phenomenology captured by contrast cases may supervene on differences in doxastic processes which manipulate the outputs of those processes without altering them.

Athanassios Raftopoulos (Chapter 11, ‘Cognitive Penetrability and Consciousness’) analyzes theories of phenomenal awareness and conscious access awareness in order to clarify the nature of the contents of late vision. He argues that studies of the attentional modulation of perceptual processing (in particular Carrasco et al. 2004; Epstein and Brooña 1986) show that late vision is cognitively penetrable. But he also claims that this cognitive influence on late vision alters the visual phenomenology of perception, with the result that scenes look different before and after the onset of cognitively driven attention. According to Raftopoulos, most accounts of perceptual awareness preclude such awareness of the visual states which result from the penetration of late vision, because on those accounts, concepts do not alter the visual phenomenology of perceptual states but instead figure alongside visual information in hybrid conceptual and visual states. Raftopoulos argues that this applies to Burge’s (2010) concept of phenomenal consciousness, Dretske’s (1993) concept of thing awareness, Tye’s (2009) account of the phenomenal character of perceptual experience, and Jackendoff’s (1989) distinction between visual phenomenology and visual understanding. On one possible construal, Block’s (2007) distinction between phenomenal awareness and conscious access awareness could accommodate the form of awareness that Raftopoulos describes, although Block does not explicitly identify such a form of perceptual awareness. Raftopoulos calls the states of perceptual awareness resulting from the cognitive penetration of late vision ‘conceptually modulated visual awareness’. Such cognitive penetration of late vision means that the content of perception is

2006; Prinz 2010). If that is so, then attention must intervene in early vision, potentially modulating it. On the other hand, there is now also empirical evidence to suggest the contrary (van Boxtel et al. 2010; Koch and Tsuchiya 2006; Lamme 2000; 2003; 2005).

These questions about awareness and attention directly affect the outcome of the debate outlined in Section 4.1 about the contents of perception, since those contents are typically thought to be divulged phenomenally. Pylyshyn (2003; 1999) is sceptical about how well introspection can inform us of the contents of different levels of perception, claiming that we introspectively confuse the panoramic conception of the visual scene which is the result of late vision with earlier stages. Raftopoulos, applying Lamme’s distinction between phenomenal and report awareness, holds that ‘the phenomenal content of perception differs from the content of the same experiences when consciously accessed’ (2009: 163). The contrast methodology, used to settle disputes based on introspective reports, may not be in a position to overcome the host of objections raised by Pylyshyn and may not curbs finely as accounts based on the neural correlates of different levels of awareness. For instance, a case has been made that phenomenal awareness, as understood by Block (1996), is the output of local recurrent processing (described in 2.2). Unless neuroscientific findings show that such processing can interact with long-term memory, the differences in phenomenology captured by contrast cases may supervene on differences in doxastic processes which manipulate the outputs of those processes without altering them.

Athanassios Raftopoulos (Chapter 11, ‘Cognitive Penetrability and Consciousness’) analyzes theories of phenomenal awareness and conscious access awareness in order to clarify the nature of the contents of late vision. He argues that studies of the attentional modulation of perceptual processing (in particular Carrasco et al. 2004; Epstein and Brooña 1986) show that late vision is cognitively penetrable. But he also claims that this cognitive influence on late vision alters the visual phenomenology of perception, with the result that scenes look different before and after the onset of cognitively driven attention. According to Raftopoulos, most accounts of perceptual awareness preclude such awareness of the visual states which result from the penetration of late vision, because on those accounts, concepts do not alter the visual phenomenology of perceptual states but instead figure alongside visual information in hybrid conceptual and visual states. Raftopoulos argues that this applies to Burge’s (2010) concept of phenomenal consciousness, Dretske’s (1993) concept of thing awareness, Tye’s (2009) account of the phenomenal character of perceptual experience, and Jackendoff’s (1989) distinction between visual phenomenology and visual understanding. On one possible construal, Block’s (2007) distinction between phenomenal awareness and conscious access awareness could accommodate the form of awareness that Raftopoulos describes, although Block does not explicitly identify such a form of perceptual awareness. Raftopoulos calls the states of perceptual awareness resulting from the cognitive penetration of late vision ‘conceptually modulated visual awareness’. Such cognitive penetration of late vision means that the content of perception is
not low-level content, and that the resulting content and phenomenology of perception are not due to penetration of early vision. On the other hand, it also means that the cognitive impenetrability of early vision does not prevent the final outputs of perception, which are accessible to the subject, from being the outputs of cognitively penetrated states. This shifts the interest of the penetrability debate to the epistemological consequences of what happens in late vision.

4.5 Representationalism and realism

In addition to his arguments against modularity (outlined in 1.3), Churchland (1988) challenged Fodor’s use of modularity to counter antirealism in the philosophy of science, arguing that a modular account of perception would not secure the neutrality of perception sought by Fodor (1988) but a form of ‘universal dogmatism’ (1988: 170). Churchland’s criticism of modularity and his criticism of the idea of using modularity to argue against antirealism are distinct: he holds that even if perception was modular, it would still not secure any form of realism.

This fundamental question is the subject of Jonathan Lowe’s contribution (Chapter 14, ‘Perceptual Content, Cognitive Penetrability, and Realism’). Lowe does not see questions of realism as independent of theories of perception, but he holds that the argument from perception to realism cannot begin from a representationalist theory of perception. For instance, he holds that it would be circular to appeal to objects and three-dimensional space prior to making an abductive inference to them, and that the argument for realism cannot proceed by ‘appeal to antecedently assumed empirical scientific theories’. Therefore, the cognitive impenetrability hypothesis is disqualified from being the starting point from which to defend realism: it is both an empirical scientific hypothesis and a representationalist theory of perception. Nonconceptualist representationalist theories like those of Dretske and of Tye, and Raftopoulos (2009), which is the target of Lowe’s discussion, are also disqualified from playing this role. However, those theories have another role to play in the defence of realism, as Lowe later explains.

Lowe claims that there is a non-representational, ‘introspectively accessible non-conceptual content of perception’ which is topologically two-dimensional—despite our ability to also experience that space as three-dimensional. We cannot explain this two-dimensional content physiologically (from the shape of retina) to argue against antirealism for the reasons already explained, but ‘the phenomenological fact itself […] is unassailable, since it is ascertainable by introspection’. From this datum we can abductively infer the existence of three-dimensional objects. By using the two-dimensional phenomena as the first step in the defence of realism, we avoid the criticism that we have inferred the world of three-dimensional objects from our subjective representations of the world. According to Lowe, representationalists have no ground from which to answer this criticism. However, representationalist theories based on empirical hypotheses—such as nonconceptualist representationalism and Fodor’s modularity hypothesis—do have an important role to play in the framework
of inferences to the best explanation: ‘the realist’s hypothesis has a genuine and substantive explanatory role to play in accounting for the structure of our perceptual experience and can be shown to fulfill this role without begging any question against the antirealist.’

In addition to the foundational issue of how to defend realism by using the cognitive impenetrability hypothesis, Lowe makes a substantive contribution to the details of the penetrability debate. He argues that the level of perceptual content isolated by the most recent version of the impenetrability hypothesis (Raftopoulos 2009) cannot play the role just outlined for representationalist theories. According to Lowe, that content does not suffice to individuate the objects of perception but only to discriminate their spatial boundaries. For example, Tibbles the cat shares spatial boundaries with a hunk of matter, but the matter can change while Tibbles remains the same (see Lowe 2007). Lowe calls hunks of matter ‘Spelke-objects’ after Spelke (1990), and argues that they are only one among many sorts of objects that perception primitively distinguishes. He argues that identity criteria for objects have to be relativized to broad, primitive sortals (like animal, hunk of matter, material artefact); these sortal-relative identity criteria are of a very broad kind and could thus be well attuned to real distinctions in nature.

Direct realism is committed to the view that subjects are perceptually aware of objects and of some of their properties directly, not by first being aware of representations of the objects and properties. In that case, the contents of perception have to be object-centred, 3D visual representations of distal objects. This view can be broken down into two claims about vision.

First, perceptual contents include expectations about how shape changes as we move relative to objects. On standard accounts of vision, such expectations require long-term memories and specifically semantic information to be applied to objects. On one hand, this suggests that direct realist accounts of perception require cognitive penetration of perception; sensorimotor knowledge without semantic information is unlikely to be adequate to produce the relevant expectations. (A suitable form of penetration may be described by Mahon and Wu (Chapter 8), who hold that functional concepts for artefacts can affect processing in the dorsal stream.) On the other hand, it means that direct realists have to explain which status they confer on the experience of objects as described by both nonconceptualists (egocentric volumetric representations, according to Evans 1982 and Peacocke 1992) and theories of vision (viewer-centred representations). If the first object representations of which we are aware are anything like those described by current theories of phenomenal awareness (see 2.2), then 3D distal objects are represented only after we have an initial egocentric visual experience of the spatial layout of the scene; in other words, they are not represented directly but indirectly. Here, the threat to naive realism is from the earliness of awareness (the awareness of stages of visual processing prior to 3D representations).

Secondly, direct realists hold that outline shapes are literally objects of vision in the world (Noe 2004: 83). This claim can be doubted on the grounds that such content
occurs too early for us to be aware of it, by using theories of vision, nonconceptual content, and awareness. For example, Briscoe (2008) has argued that conscious representations of outlines are likely to be cognitively driven achievements.

Costas Pagondiotis (Chapter 15, ‘Cognitive (Im)penetrability of Vision’) asks whether the portion of perception defined as impenetrable by Pylyshyn allows perception to play the role of anchoring the subject to the world. He argues that in this respect, the cognitive impenetrability hypothesis has led to a philosophically undesirable dilemma: either vision is not dependent on cognition and consequently vision is alienated and disconnected from thought, or else vision is dependent upon cognition, with the result that the distinction between vision and thought is blurred.

He suggests two ways out of this dilemma. The first remains within the framework of the representationalist–computationalist theory of mind. It involves defining the cognitive penetrability of early vision as its dependence on other systems which involve either non-propositional representations or propositional but domain-specific representations. In either case, vision would not be ‘promiscuous’ like thought is, so this could provide the distinction between vision and thought which, according to Pagondiotis, is not provided by Pylyshyn’s account. The second way out of the dilemma, favoured by Pagondiotis, draws on theories of perception that reject the representationalist–computationalist theory of mind and support an anti-intellectualist approach to perception, stressing the dependence of perception on practical non-propositional knowledge (Dreyfus 1992; Haugeland 1998; Hurley 1998; Noë 2004; O’Regan 2011). Pagondiotis argues that vision is directly penetrated by practical non-propositional knowledge, and that as a result, visual content is object-dependent and inextricably connected to an embodied agent, instead of being an isolated atomic given disconnected from the perceiver.

According to Pagondiotis, taking visual content to be penetrated by practical non-propositional knowledge paves the way for a defence of the view that perception brings us into direct contact with the world, and for a form of realism based on the attunement between perceptual content and world. That attunement is exemplified by the dependence of perceptual content on sensorimotor knowledge and anticipation. While Pagondiotis sees practical non-inferential knowledge as a form of Churchland’s diachronic penetration, he holds that perception’s dependence on such knowledge is rational without being either inferential or propositional. On this basis, he argues that practical non-propositional knowledge does not constitute a form of theory-ladenness and provides ‘a sufficient ground for scientists and ordinary observers to communicate and to find out what is or is not veridical’—satisfying the epistemological objective set for modularity by Fodor (see 1.3).

### 4.6 Action

Apart from the actions involved in interactive accounts of perception (see previous section, and Pagondiotis, Chapter 15), hypotheses about cognitive penetrability and
impenetrability are relevant to theories in the philosophy of action that distinguish different levels of action like distal, proximal and motor intentions (Pacherie 2008), or prior intentions and intentions-in-action; and to accounts of how actions inherit the goals of prior intentions, such as Israel et al.’s (1993) concept of ‘belief-how’.

Pylyshyn (2003) has argued that motor systems are cognitively impenetrable, citing evidence about the dorsal ‘vision-for-action’ stream (Milner and Goodale 1995; Ungerleider and Mishkin 1982) and findings which suggest its autonomy from the processes that subserve object recognition (2003: 152). Evidence that the final stages of vision and first stages of motor control share information in the same format, and that conscious planning impairs detection and response (Hommel 2004), suggest direct interactions between dorsal system and motor systems without cognitive mediation; and much recent work in psychology shows direct ties between perception and action (Beets et al. 2010; Humphreys and Riddoch 2001; 2007; Hommel 2004; Klatzky et al. 2010; Goodale and Milner 2004; Riddoch et al. 2003).

However, Briscoe (Chapter 7) argues that visual processes which compute size for perception are penetrated by distal intentions. Vishton et al. (2007) not only found that grasping affects the visual experience of size, but that even ‘listening to a description of a reaching task’ affected size perception. As Briscoe points out, this implies that forming the intention to act suffices to influence size perception. If he is right, then there are cases in which action determines perceptual content in a much more intimate way than is usually thought to occur, for instance by sensory exploration. One could think of sensory exploration as an activity in which we probe the visual scene by directing visual resources and letting them process the scene visually, but that clearly does not apply to the case just described.

Brad Mahon and Wayne Wu (Chapter 8, ‘Cognitive Penetration of the Dorsal Visual Stream?’) argue that the contents of intentions, semantic information about functional kinds, or both, penetrate the computations of motor response in the dorsal visual stream. They defend this thesis in a framework which sees action as a problem to be solved by the brain: intentions represent actions, actions are specific input–output mappings, and the brain must produce the mappings. The problem can be split into components on the basis of dissociations between reaching to grasp (impaired in optic ataxia) and object use (impaired in apraxia). Reaching can be determined by the dorsal stream’s visual computations, but information about which manipulations are appropriate for object use constitute stored knowledge; they are comparable to lexical representations of words (Rothi et al. 1991), and count as semantic information. Thus, prima facie, the dorsal system cannot solve the entire problem of producing an action plan. (One solution could be provided by visual-motor memories—Allport 1985; Gallese and Lakoff 2005. Mahon and Wu argue against this solution by adducing evidence from object-use apraxia and optic ataxia.)

Mahon and Wu show that the action is highly unlikely to be produced by an influence of semantic information on the inputs to, or outputs from, the dorsal-stream processing. For instance, if the ventral stream analyzed the properties of artefacts
and highlighted their relevant parts for action, that would still underdetermine the action plan. At output level, it could be thought that the dorsal stream provides a set of action plans and then one is selected. What militates against this proposal is that dorsal computations are thought to be restricted to immediate manipulation, while the grasping component of reach-to-grasp movements anticipates manipulation. Thus, patients with object-form agnosia are unimpaired for reaching but impaired for grasping objects in the right way to perform the appropriate action. They are also impaired for integrating information about the surface properties of an object, such as whether the handle was slippery, into their initial grasp.

This suggests that the action plan cannot be determined by the dorsal stream without help from semantic information about functional kinds and other properties of the object to be acted upon. Mahon and Wu consider two possible sources of such semantic information. One is that representations of objects in the contents of intentions to act (on those objects) directly transfer information for computation to the dorsal stream. For example, representing how one intends to use a hammer will determine how one reaches to grasp it. Alternatively, functional concepts for objects may transfer information over which the ventral stream computes; this is consistent with deficits in semantic dementia, which affects cognition of the function of artifacts.

### 4.7 Perceptual warrant

According to Siegel (2012), cognitive penetration that determined which high-level properties are represented in perception would vitiate the perceptual justification of belief: if Jill visually experiences Jack’s face as angry because she believes that Jack is angry, then visually experiencing Jack’s face as angry can no longer justify her belief that Jack is angry. Such epistemically illicit feedback from beliefs to perception would threaten the internalist principle of justification that Lyons (2011) calls ‘seemings internalism’: ‘one’s belief that \( x \) is \( F \) is prima facie justified any time it is based on one’s having a perceptual experience that \( x \) is \( F \).’

The dependence of perceptual experience on conceptual states that Siegel has in mind is meant to apply even when eye fixation and attention are constant, so the attention-shift argument cannot be used to deny that it is cognitive penetration specifically (see Section 3) that damages perceptual justification. Alternatively, one might seek to deny the inference from cognitive penetration to vitiation of justification: the threat to justification may not materialize if the effect of semantic information is restricted to cognitive phenomenology, understood as distinct from the sensory phenomenology to which the epistemic subject appeals during internal justification. However, this route is also blocked because Siegel (2010) holds that the representation of high-level properties affects sensory phenomenology specifically.

A way to deal with potential threats to perceptual warrant from cognitive penetration is proposed by Jack C. Lyons (Chapter 3, ‘Unencapsulated Modules and Perceptual Judgment’). As explained in Section 2.1, Lyons holds that information-processing systems can be modular even if they are not informationally encapsulated,
including from cognitive processes. But he also denies that penetrability would necessarily damage perceptual warrant by arguing that cognitive penetrability does not diminish the overall reliability of perception. Justificatory circularity would require a basing relation, not just a causal relation, to obtain between penetrating and penetrated states and between perceptions and resulting beliefs. In Siegel’s example, the content of the perception is not based on long-term memories, although it is caused by them, and the content of the belief is not based on the content of the perceptual experience although it may be caused by it.

Lyons (Chapter 3, and 2011) also argues that penetrability can be an epistemologically benign influence in a reliabilist framework, for instance when it enhances the rapid classification of stimuli. If one is on a hike and watching out for snakes, cognitive penetration may allow such rapid classification and even make certain stimuli more salient. Another example is perceptual learning (if construed as a form of ‘diachronic penetration’; see 2.4): as an improvement of perceptual discernment in a specific domain, it appears to be an epistemologically benign influence on perception. The association of impenetrability with epistemological virtuousness can also be doubted. Lyons (2011) considers three possible explanations of a case-type of cognitive penetrability, the penetration of colour perception:

1. subjects’ current, occurrent belief that bananas are yellow has genuinely top-down effects on the perceptual state, making the banana look more yellow;
2. subjects’ longstanding belief that bananas are yellow has the top-down effect; and
3. subjects’ history with yellow bananas has produced an associative connection whereby low-level perceptual features of bananas (and not the banana identification itself) prime yellow in the colour detection system, producing a lateral, rather than top-down, effect on the perceptual state.

Explanation (2) amounts to cognitive penetration, while (3) does not (see Section 2.1). Yet there is no difference in the epistemological value of (2) and (3) in a reliabilist framework according to Lyons—despite the fact that (3) is currently used in the literature as a denial of cognitive penetration, while (2) is considered an admission of cognitive penetration. Moreover, (1) is epistemologically more pernicious than (2)—a difference that does not show up on existing definitions of cognitive penetrability.

References


