nels to optimize the encoding of information (e.g., contrast or orientation) in the task considered. Further studies are currently carried out to understand the parameters of spatial filtering that are under such cognitive influence.

Generalizing from these results, it would seem that the "allocation of attention to visual properties prior to the operation of early vision" does not reduce cognitive penetrability to a few situations of recognition. Instead, the most common situations concern the basic (*bird, car*) vs. subordinate (*sparrow, Mercedes*) categorizations of objects which are known to require different cues from the visual input (as LSF and HSF cues were best-suited to different face categorizations in the experiments just described). It is therefore conceivable that the common basic versus subordinate categorizations of identical objects would elicit distinct perceptual settings of early vision.

To conclude, I presented evidence of cognitive penetrability which support Pylyshyn's criterion. I then argued that the supposedly restrictive criterion of penetrability could in fact cover many situations of recognition. Two main issues are now in the empirical arena: (1) Is cognitive penetrability the exception or the rule of visual categorization? and (2) What parameters of early visual functions are under cognitive influence? Cognitive penetrability should not be settled *a priori*.

ACKNOWLEDGMENTS

The author wishes to thank Pascal Mammassian for helpful comments. The described research was partially funded by ESRC grant R000237412 awarded to Phillipe G. Schyns and a Fyssen Foundation fellowship awarded to Aude Oliva.

Perception, inference, and the veridicality of natural constraints

Manish Singh[a] and Donald D. Hoffman[b]

[a]*Perceptual Science Group, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139;* [b]*Department of Cognitive Sciences, University of California, Irvine, CA 92697-5100.*
manish@psyche.mit.edu **www-bcs.mit.edu/~manish**
www.socsci.uci.edu/cogsci/personnel/hoffman/hoffman.html
ddhoff@uci.edu

Abstract: Pylyshyn's target article argues that perception is not inferential, but this is true only under a narrow construal of inference. A more general construal is possible, and has been used to provide formal theories of many visual capacities. This approach also makes clear that the evolution of natural constraints need not converge to the "veridical" state of the world.

Pylyshyn's target article distinguishes the natural constraints approach to vision from the inference approach. This distinction is real if one thinks of inference, as Pylyshyn does, as being necessarily deliberate and conscious, or at least as having unrestricted access to the individual's beliefs and goals. But clearly it is possible to think of inference more generally – e.g., as a mapping from premises to conclusions that preserves some relevant structure. Under such a formal definition, perception is very much an act of inference. Indeed, it is becoming increasingly popular to think of perception as an inductive inference (i.e., as an inference whose premises do not logically force its conclusions), that is appropriately modeled using Bayesian statistics (e.g., Bennett et al. 1996; Freeman 1996; Knill & Richards 1996). Such an approach has been used profitably to provide formal theories of many visual capacities (Brainard & Freeman 1997; Bulthoff & Yuilee 1991; Freeman 1994; Richards et al. 1996; Singh & Hoffman 1998). It has the advantage of allowing one to study perceptual inference and higher-level, cognitive (or "rational") inference within a common framework. Furthermore, it does not violate the cognitive impenetrability thesis because the premises for perceptual inference need not include the individual's beliefs and goals.

Such an inference based approach to perception also clarifies the role of natural selection in the evolution of natural constraints. For example, the target article tacitly assumes that, over the course of evolution, the constraints that human vision uses, and therefore its representations of the visual environment, become more and more veridical. The truth of this claim depends critically on how we understand the term "veridical." If by veridical we mean a resemblance relation – that, somehow, our visual representations begin to resemble the objective world – this is more than we can safely claim. We have no way to step outside ourselves – outside our perceptual constraints and representations – and evaluate to what extent these representations resemble the objective world. Every percept, observation, and judgment on our part is a conclusion (in the general sense of *inference* mentioned above) based on our current constraints and knowledge. In the language of Bayes,

$$P(S \mid I) = P(I \mid S)\, P(S) / P(I)$$

where I is an image, or set of images, presented to the visual system (such the P(I) is not zero), and S is a possible scene representation of the image. $P(S \mid I)$ is the posterior probability of the scene S, given the image I. $P(I \mid S)$ is the likelihood of obtaining the image I, given that the scene is S. And P(S) is the prior probability of S, reflecting the constraint embodied by the observer. This formalism makes clear that all we ever see and experience are our posteriors – based on our current priors. Over the course of evolution, our current posteriors may become our future priors. But this recursive updating by no means guarantees that the priors will eventually converge to the true probability measure that defines the world – which is precisely what would be required in order to claim that the natural constraints embodied in our visual system, and the resulting visual representations, eventually resemble the objective world. Examples such as the Ames trapezoidal window with the rod placed inside, the Pulfrich double pendulum (sect. 5.2, last para.), and other physically impossible percepts (see Fig. 1) certainly make a strong case for the cognitive impenetrability of perception. But they also underscore the lack of any resemblance relationship between our perceptual representations and the objective world.

However, a weaker interpretation of the term "veridical" is possible – one that is based not on resemblance, but on utility. In other words, the natural constraints that evolve do so because they endow the organism with a survival advantage, irrespective of whether the resulting representations bear a resemblance relationship to the objective world. Cockroaches are quite successful evolutionarily, for example, and this must be attributed to useful and sophisticated representations that they have developed. This does not mean, however, that these representations have con-

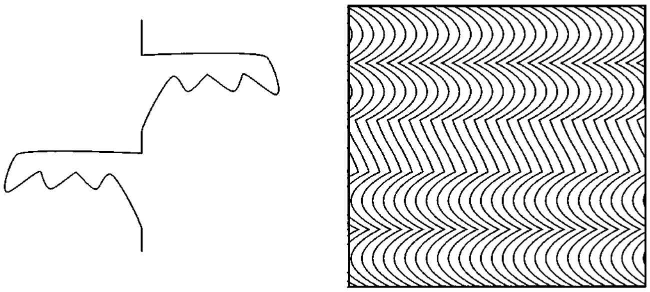

Figure 1 (Singh & Hoffman). In order to have a consistent physical interpretation, a plane curve must enclose physical material consistently on one side. Local geometric constraints can, however, override this requirement, leading to globally inconsistent figure-ground assignments (see Hoffman & Singh 1997). Such physically impossible percepts underscore the lack of any necessary resemblance relationship between our perceptual representations and the objective world.

verged or are converging, to the true state of the world. A simple way to clarify the difference between utility and resemblance is to think of an icon interface on a personal computer (Hoffman 1998). The folder, file, and trash icons on the computer screen provide a convenient, faithful, and useful way to interact with the diodes and transistors that form the circuitry of the computer. It would be a mistake, however, to claim that the shapes and colors of these icons bear any resemblance relation to the circuitry.

Furthermore, it is not sufficient to say that although perceptual representations do not capture all dimensions of the true state of the world, they do veridically represent (in a resemblance fashion) some of these dimensions – or at least they will do so eventually. It is also possible, for example, that the perceptual constraints *impose* certain dimensions on incoming data, which have no canonical dimensional structure to begin with. To take a toy example, let us say we have a "world" of words – a set with no canonical dimensions. Different organisms with different perceptual constraints could impose different dimensional structures on this set. For example, one organism could structure the set of words according to the alphabetical position of their last letter (dimension 1), and how close their first letter is to the letter "m" in the alphabet (dimension 2). Another organism might impose a very different structure on this set (based, perhaps, on the syntactical category of a word). The imposed structure in each case might, perhaps, enable the organism to interact usefully with its environment of words, but this would not mean that the organism has discovered the "true" dimensional structure of its environment – because there wasn't one to begin with.

Expert perceivers and perceptual learning

Paul T. Sowden

Department of Psychology, University of Surrey, Guildford, GU2 5XH, United Kingdom. **p.sowden@surrey.ac.uk**

Abstract: Expert perceivers may learn more than just where to apply visual processing, or which part of the output from the visual system to attend to. Their early visual system may be modified, as a result of their specific needs, through a process of early visual learning. We argue that this is, in effect, a form of long-term, indirect cognitive penetration of early vision.

Are early vision and visual perception the same? We start on a slightly pedantic point. The definition of visual perception (VP) is at the heart of the problem which Pylyshyn's target article addresses, and it is the very issue of its continuity or discontinuity with cognition that has caused such definitional problems. Many authors would include the cognitively mediated direction of attention and recognition processes, which clearly affect the phenomenal, conscious content of perception (e.g., sect. 6.1), as being part of VP. Pylyshyn does not attempt a definition directly, but argues for the existence of a distinct module which he terms early vision (EV) that is impervious to cognitive influences. He does, however, appear to acknowledge its distinction from VP, stating

> we will conclude that although what is commonly referred to as "visual perception" is potentially determined by the entire cognitive system, there is an important part of this process – which, following roughly the terminology introduced by Marr (1982), we will call early vision – that is impervious to cognitive influences (sect. 1, para. 2)

Confusingly a distinction is not adhered to or consistently brought out throughout the target article. In places, arguments are made which might support the view that EV is cognitively impenetrable, but which are presented as showing that *VP* is cognitively impenetrable. Without a clear position on the definition of VP, we feel that the case for the impenetrability of VP cannot be established. However, the case for the impenetrability of EV is argued convincingly; nevertheless, we next argue for an exception to the impenetrability of EV.

Early visual learning. The flexibility of our perceptual systems is important if we are to maximise the effectiveness with which we can interact with our environment. It has been argued that EV has been modified phylogenetically to match species specific needs by embodying a form of knowledge within the visual system (e.g., Pylyshyn's "natural constraints" sects. 5, 5.1; Barlow 1997). We believe that EV can also be modified ontogenetically, through a type of perceptual learning (Karni & Bertini 1997) that here we call early visual learning (EVL), to meet the needs of a specific individual. This process may be mediated, as Pylyshyn suggests (sects. 6.3, 6.4), by the direction of attention to the relevant stimulus properties. There is increasing evidence of the importance of attention to the relevant features in order for learning to occur (e.g., Ahissar & Hochstein 1993; Shiu & Pashler 1992) and in fact in some cases apparent EVL may be better explained as resulting from the operation of selective attention mechanisms (O'Toole & Kersten 1992; Sowden et al. 1996). Our own work suggests that expert perceivers (e.g., radiologists) learn which dimensions of visual analysis to attend to, and that, as a consequence, their analysis of them becomes enhanced through modifications *within EV*[1] (see Davies et al. 1994; Sowden et al., submitted). Further, we have shown that even language may shape basic perceptual sensitivity to a small extent, through our work on colour perception (see Davies et al. 1998), perhaps by operating to influence the direction of attention to linguistically salient dimensions of visual analysis, which then encourages subsequent EVL.

Thus far, the work described does not necessarily contradict Pylyshyn's view. As a result of acquiring knowledge, an individual modifies the dimensions of visual analysis to which they attend in accord with their particular needs. Then through a process of EVL modifications to the operation of EV take place.[2] However, whilst we agree this process may involve "shaping basic sensors, say by attenuating or enhancing the output of certain feature detectors (perhaps through focal attention)," we do not agree that this does not "alter the contents of perceptions in a way that is logically connected to the contents of beliefs, expectations, values, and so on" (sect. 1.1, para. 1) and which Pylyshyn consequently argues does not count as cognitive penetration. These modifications can clearly be connected to the individual's expectations, values, and so on (albeit mediated by attention), as shown by the learning of expert perceivers (see previous para.). Consequently, such learning could be said to indicate a form of indirect cognitive penetration. We feel it is worth describing this type of learning as a special form of cognitive penetration because such a description implies the long-term adaptiveness of the visual system to an individual's knowledge and needs.

In addition, it is not clear that attention can be afforded the degree of separation from EV that Pylyshyn suggests, because the activity of neurons in both striate and prestriate visual cortex appears to be moderated by attention (e.g., Motter 1993). Conceivably, the route to modification by EV may not be quite so indirect as we have suggested. On this point we await resolution of the degree to which attention is functionally and structurally separate from sensory processing.

So far we have argued for a special case of indirect cognitive penetration which operates on a relatively long-term time scale, where clearly an individual's cognitions have no immediate impact on their perception. However, we would like to finish on another point related to the performance of expert perceivers. Pylyshyn considers the case of chick sexing (Biederman & Shiffrar 1987) and argues that instruction teaches the sexer how "to bring the independent visual system to bear at the right spatial location" (sect. 6.2, para. 4). An alternative explanation is that the effect of instruction is to encourage the formation of a new "figure" which facilitates discrimination between the sex of the chicks. This could be analogous to the functional feature creation of Schyns and Rodet (1997) and Schyns et al. (1998) and, given the rapid time scale on which learning occurs in the case of chick sexing, may indicate that this process is not one of "tuning of basic sensory sensitivity by task specific repetition" (sect. 6.3, para. 3), but in fact