# Against a "Mindless" Account of Perceptual Expertise

Abstract:

According to Hubert Dreyfus's famous claim that expertise is fundamentally "mindless," experts in any domain perform most effectively when their activity is automatic and unmediated by concepts or cognitive processes like attention and memory. While several scholars have recently challenged the plausibility of Dreyfus's "mindless" account of expertise for explaining a wide range of expert activities, there has been little consideration of the one form of expertise which might be most amenable to Dreyfus's account – namely, perceptual expertise. Indeed, Dreyfus's account of expert coping is ultimately an account of perceptual expertise, in that an expert's intuitive situational responses are thought to rely on a sophisticated repertoire of perceptual skills. In this paper, I examine the feedforward model of sensory processing that Dreyfus uses to illustrate the perceptual underpinnings of expert action, and consider its resonance with psychological research that characterizes perceptual expertise as being automatic, holistic, pre-attentive, and non-cognitive in nature. However, citing competing empirical research, I argue instead that Dreyfus's model of perceptual expertise cannot adequately explain the integral roles of attention, memory, and conceptual knowledge in expert object recognition. I conclude that the Dreyfusian model of perceptual expertise fails – the perceptual repertoire of skills that grounds expert object recognition is not operative in isolation from the expert's conceptual repertoire.


Keywords: Hubert Dreyfus; Perceptual Expertise; Object Recognition; Attention; Perceptual Concepts; Non-Conceptualism

## Hubert Dreyfus on Perceptual Expertise: A Critique

## 1. Introduction: Dreyfus on Memory, Attention, and Expert Intuition

According to Hubert Dreyfus's (2005/2007a/2013) influential account of expertise, an expert skillfully copes with the environment in a fundamentally "mindless" manner. For Dreyfus, activity reaches the level of expertise when its optimal performance becomes automatic and unguided by conscious thought. In the flow of absorbed, unthinking activity, experts directly perceive and respond to the world without the mediation of concepts or propositionally structured mental representations. Indeed, Dreyfus claims that the experience of absorbed expert coping is essentially incapable of being expressed in propositional terms: To mentally represent the world as a set of facts to be known is to open a gap between the mind and the world that does not exist when we are mindlessly immersed in an environment of attractive and repulsive affordances.

Dreyfus sees this gap between mind and world as being initially opened through the operation of cognitive factors like memory and attention. These factors not only disrupt absorbed expert coping, but also transform coping experience into the content of conceptual judgments and detached rational reflection. Apparently taking cognitive/declarative memory as his target, Dreyfus is skeptical that this type of memory can play any role in what he calls "intuition," or an expert's holistic ability for immediately "zeroing in" on whichever possible responses to a situation are most optimal (see Dreyfus 1972: 18 & passim). In contrast to artificial intelligence programs, human experts don't rely on the brute-force processing of innumerable and discrete pieces of information stored in memory in order to determine an optimal set of possible response. Instead, they rapidly zero in on the most appropriate response by perceptually recognizing the similarities between the current situation and previously experienced situation. With the achievement of full mastery, the expert acquires a sensory sensitivity to subtle perceptual patterns that allows her to circumvent any deliberative process. Seeing a current situation to be relevantly similar to a past situation, the expert then automatically knows how to respond.

The perceptual patterns which experts ultimately learn to recognize are holistic in nature. That is, expert intuition processes a situation as a sort of gestalt whose response-relevant significance cannot be decomposed into constituent parts or patterns. Experts do not judge a situation and determine what to do by classifying and associating salient parts in accordance with some stored mental rule. As a result, Dreyfus is skeptical of theoretical accounts which claim that experts recognize situations by drawing upon the memory of basic chunks, or typical groupings of features that are further associated with condition-action rules stating that if the grouping is present, then a certain response is optimal. Because Dreyfus links the reliance on these simple conceptual rules with being a non-expert, he claims that the expert must be recognizing the situation as a whole, rather than as made up of component chunks (Dreyfus and Dreyfus 1988: 34).[1]

---

1  For discussion of how recent developments of chunking theory respond to Dreyfus's objections, see Gobet and Chassy 2009.

Whether as chunks or as holistic patterns, Dreyfus argues, the mnemonic retrieval of mental representations cannot ground the expert's ability for recognizing an innumerable number of situations on the basis of past experiences. For one, there is no evidence within the phenomenology of absorbed coping for such a retrieval process; nor would the rapid performance of skillful action allow any time for retrieval to take place (Dreyfus and Dreyfus 1988: 28). Instead of memory storage and retrieval, Dreyfus (2002) prefers to speak of past experiences as reshaping the connections between nodes in an expert's feedforward neural network. This move further divorces expert intuition from the processing of discrete mnemonic representations with conceptual content. Representations of past experiences are not first stored and then consciously recalled by experts as they are rapidly responding to novel environmental solicitations. Rather, past experience conditions the expert's response to new input at a sub-representational, non-cognitive level. Hence, the expert's repertoire of holistic, unchunked situations is stored not as memory representations in the mind, but as bodily dispositions for directly responding to perceived situations (Ibid.: 374).

Dreyfus is additionally led by the automatic and non-cognitive nature of expert intuition to conclude that experts must not pay attention to what they are doing, or to the objects of their activity (2007b: 374). He seems to view the exercise of attention as a type of conceptual capacity that involves self-consciously experiencing oneself as a monitoring subject or ego. Such a self-conscious stance is only taken when something disrupts one's intuitive responses, and one consequently has to deliberate as a non-expert would about what rules for action one should follow (2007a: 361; 2002: 369). Within the flow of absorbed coping, however, there is not even a trace of the "I," nor any minimal awareness of oneself (2007b: 373-4). Moreover, beyond the disruption of absorbed coping, the deployment of attention also brings about a "radical transformation" of absorbed experience and its content. Since Dreyfus considers attention to be a conceptual capacity, he takes the content of attentive experience to be propositionally structured, and hence to be essentially different from the non-propositional character of absorbed experience. Dreyfus asserts that it is only when attention is directed to the affordances present in absorbed coping that we can then experience a world of stable objects with abiding properties, which are the sorts of objects about which we can rationally form propositional beliefs, judgments, and inferences. Attention thereby conceals the level of non-conceptual perception and coping at which the world is primordially given (Dreyfus 2005: 61; 2007a: 363).

Now, several scholars have questioned the plausibility of Dreyfus's "mindless" model of expertise for explaining a wide range of expert activities, arguing that Dreyfus has oversold the automatic character of expertise and the extent to which absorbed coping is incompatible with conscious and cognitive control (Montero 2016; Fridland 2017; Sutton et al. 2011). Dreyfus's specific claim that optimal expert performance occurs in the absence of memory and attention can be similarly disputed. Expertise in a diverse variety of fields has been shown to rely on the formation of chunks, i.e., meaningful information-patterns encoded in, and retrievable from, long-term memory (Feltovich, Prietula, and Ericsson 2006). And despite his exhortation that experts must not attend to their activity if they are to remain in an optimal state of flow, paying attention is not inherently incompatible with an expert's automatic and intuitive coping, and can actually facilitate the flexible adaptation of otherwise automatic responses to contextual demands

(Geeves et al. 2014; Toner, Montero, and Moran 2015; Christensen, Sutton, and McIlwain 2016). To Dreyfus's claim that attending to objects interferes with an expert's response to them, a wealth of research shows just the opposite – both expert and novice athletic performance is improved by adopting an external focus of attention, or in certain cases to internal kinesthetic cues (Wulf 2013; Christensen, Sutton, and McIlwain 2015; Toner, Montero, and Moran 2016: 309-10; Breivik 2013: 101; Montero 2016: 179-82). The larger point is that expertise entails the presence of refined attentional skills, rather than the lack of any attention at all.

Yet, leaving aside the above challenges to Dreyfus's "mindless" model, there has been little consideration of the one form of expertise which might be most amenable to Dreyfus's account – namely, perceptual expertise. In fact, Dreyfus's general account of expert coping ultimately rests upon an account of perceptual expertise, in that an expert's intuitive situational responses rely upon a sophisticated repertoire of perceptual skills for discerning subtle patterns of similarity and difference. It would be true to form for Dreyfus to claim that expert object recognition not only enables the unmediated bodily responsiveness to environmental affordances which defines absorbed coping, but that object recognition is itself is a form of absorbed coping.

Thus in section 2 of this paper, I develop the plausibility of a "mindless" model of perceptual expertise by canvassing empirical support for the claim that expert perceptual processing is holistic, automatic, and pre-attentive. In section 3, I raise an initial objection to a Dreyfusian characterization of perceptual expertise as being essentially non-conceptual and non-cognitive, by showing especially that real-world domains of expert object recognition rely upon the expert's possession of conceptual knowledge. Section 4 considers how Dreyfus could respond by incorporating even this knowledge into the feedforward model of sensory processing that he uses to illustrate the perceptual/non-conceptual underpinnings of expert action. Then in section 5, I offer another refutation of the Dreyfusian account of perceptual expertise, by drawing on recent studies of perceptual expertise which support the claim that a perceiver's cognitive, personal-level states of intention and attention enable the patterns of neural activity that underlie expert object recognition. I conclude in section 6 by considering how these findings cohere with general theories of expertise that, contra Dreyfus, acknowledge the roles of attention, memory, and cognitive control in absorbed activity.

## 2. The "mindless" character of perceptual learning and expertise

Dreyfus acknowledges that becoming a master in most any domain involves the acquisition of skills for "holistic similarity recognition," while also denying that what the master has acquired are concepts. He writes that what masters learn through practice "are not critically justifiable concepts but sensitivity to subtler and subtler similarities and differences of perceptual patterns. Thus, learning changes, not the master's mind, but his world" (2013: 35). Dreyfus further states that this acquired sensitivity grants an expert a "rich *perceptual* repertoire – the ability to respond to subtle differences in the appearance of perhaps hundreds of thousands of situations – but it requires no *conceptual* repertoire at all" (2005: 58). These skills of perceptual pattern recognition are what uniquely grant experts with the intuitive ability for rapidly selecting an optimal response on the basis of perceived similarities with previously experienced situations.

We will now consider the extent to which Dreyfus's account is corroborated by psychological studies that characterize perceptual expertise as an automatic, pre-attentive, and largely sensory process.

To defend his view that experts learn to discern subtle perceptual patterns in a way that bypasses the conceptual/cognitive mind, Dreyfus might first point to the phenomenon of perceptual learning, whereby perceivers acquire enhanced abilities for performing a certain perceptual task or processing a certain stimulus merely through repeated exposure to that task or stimulus. Perceptual learning of this sort is apparent in the trained improvement on tasks such as discriminating vernier acuity, visual textures and gratings, motion direction, and stereoscopic depth (see Lu et al. 2010). Given that the acquisition and exercise of such discriminatory abilities can occur outside of the perceiver's conscious awareness, psychologists have largely defined perceptual learning in opposition to declarative learning, i.e., the process of acquiring knowledge of facts and events that can be consciously recalled and verbally described, and procedural learning, which involves consciously learning skillful patterns of action that can either be deliberately or automatically exercised. By contrast, perceptual learning can be viewed as an implicit learning process that does not leave the perceiver with any explicit, reportable sense of what has been learned (Kellman and Garrigan 2009: 55-6).

Similarly at the neural level, the improvements in perceptual discrimination induced through repeated exposure evidently leave little to no cognitive trace, instead producing long-term adaptations in the regions of the brain responsible for low-level sensory processing (Karni and Sagi 1995; Raftopoulos 2001). Enhanced task performance is additionally taken to be restricted to specific retinal locations as well as to specific stimuli. That is, a trained increase in sensitivity among neurons in one part of the visual field won't necessarily transfer to untrained neurons in another part, suggesting again that the learning effect takes place in the earliest parts of the visual cortex where neurons are still retinotopically organized (Karni and Sagi 1995: 96). Moreover, EEG evidence of the changes in performance due to perceptual learning have been detected within 100 ms after stimulus onset, likely before higher-level visual and cognitive processes would exert any top-down influence (Goldstone and Bygre 2015: 816-7). Hence, in terms amenable to Dreyfus's account, Fahle (2002: x) summarizes, "Perceptual learning leads to implicit memory, to 'knowing how,' to a 'memory without a record' and is often very specific for rather low-level attributes of the stimulus learned."

Nonetheless, it is clear that perceptual learning as a passive, stimulus-driven process is insufficient for developing the kind of expert intuition that supports absorbed coping in Dreyfus's account. Granted, there is a parallel between perceptual learning and how Dreyfusian experts acquire an increased sensitivity to perceptual patterns through repeated practice in the absence of declarative learning – but the similarity ends there. Even rudimentary forms of absorbed coping, like the act of reaching out and turning a doorknob, are too complex to be guided solely by the kinds of perceptual skills acquired through low-level perceptual learning. If perceptual categorization were restricted in the sorts of ways that low-level learning effects are purported to be, then even slight variations in the retinal location of a stimulus would prevent experts from recognizing what they are seeing. Expert intuition must instead rely on perceptual object

recognition, which necessarily abstracts away from variations in sensory features that are irrelevant to an object's category membership. For instance, an expert indoor rock climber can be presented with handholds that differ in color, shape, and size, and still rapidly intuit that they all require the same type of grip posture (see Bläsing et al. 2014). Similarly, a chess expert can intuit the correct response to a given board position equally well whether looking at a physical three-dimensional board or a two-dimensional board printed in a book. The representations formed in early vision, however, are highly sensitive to sensory variations – consequently, object recognition must draw upon the more abstract and categorical representations formed in higher-level areas of vision, rather than purely those areas affected by low-level perceptual learning.[2]

Accounting for Dreyfusian intuition, then, shifts us from low-level perceptual learning to perceptual expertise, understood as the enhanced ability to perceptually recognize and distinguish between similar instances of the same class. Examples of real-world perceptual expertise include: the radiologist's ability to diagnose a condition on the basis of subtle perceptual cues in an x-ray; the sommelier's ability to distinguish subtle tastes and odors of wine; the ability of a bird-watcher to rapidly recognize a species of a bird in a dense forest; and the musician's ability to differentiate two musical tones of similar frequency. Though these examples may give the impression that perceptual expertise is the province of highly specialized experts who have undertaken years of deliberate training, most adult humans have enough practice to be perceptual experts in at least two domains, namely face recognition and fluent reading. For all of these domains of expertise, experts will consistently outperform novices in relevant perceptual categorization tasks. Whereas a novice makes basic-level categorizations (e.g., "bird," "dog") faster than subordinate-level categorizations (e.g., "robin," "terrier"), experts can perceptually categorize objects at both levels equally rapidly (Palmeri and Gauthier 2004: 297).

The advantages that perceptual experts have over novices in perceptual categorization have been attributed in part to how experts parse visual stimuli differently than novices. These differences have been linked in part to activity in the fusiform face area (FFA), a high-level area of the visual cortex's ventral stream.[3] The FFA has been in implicated processing stimuli across a wide range of domains beyond just facial recognition – expert performance in visually recognizing cars, birds, butterflies, artificial computer-generated objects, chess positions, and x-rays all have been correlated with increased activity in FFA (see Bilalić 2016). What all these

---

2    It is further debatable whether perceptual learning is itself a purely low-level phenomenon occurring only at the earliest stages of vision – see Kellman and Garrigan 2009: 72-5 and Wang et al. 2016.

3    To clarify, the ventral visual pathway is hierarchically structured into low-level and high-level areas according to the level of abstraction and complexity at which visual stimuli are processed. The ventral stream starts in the occipital lobe, branches off from the primary or "striate" visual cortex (V1) and secondary or "prestriate" visual cortex (V2), and then proceeds through to "extrastriate" areas V3 and V4. These four visual regions are considered to be low-level areas because they are responsive to rudimentary sensory features like color and shape. An intermediate state of processing occurs at the lateral occipital complex, where sensory features are integrated into the representation of a coherent three-dimensional object. The ventral stream culminates in the inferior temporal cortex located in the temporal lobe. Here, high-level regions are responsible for identifying objects under categories (e.g., "car," "face," "dog") that abstract away from variations in low-level stimuli. Of course, there are also recurrent feedback signals from high-level areas back down to low-level areas. See Grill-Spector 2010 for a more detailed overview of the high-level and low-level visual regions involved in object processing.

types of stimuli have in common is that they've been found to be processed by experts in a holistic manner. That is, whereas novices selectively attend to a few parts of an object in order to categorize it, experts will attend to the object as an integrated whole. Members of certain categories – such as human faces or chess positions – often share a prototypical configuration of parts, so it makes sense that an enhanced ability for identifying and discriminating between category members would rely on attending to multiple parts and the configural relations between them. A novice who is looking for a certain plant in the woods, for instance, may have to categorize objects by deliberately following rules – e.g., "look for smooth, non-serrated leaves with an elongated oval shape" – and scanning parts individually. Expert categorizers, however, can employ holistic processing, which parses complex visual patterns by binding together features into larger configurations that get encoded as a single meaningful unit, in a manner akin to memory-based chunking (Goldstone and Bygre 2015: 821). Unitizing stimuli in this way facilitates expert object recognition, as it allows for an increased amount of perceptual information to be compared with category-exemplars or templates retrieved from memory, a process which is less deliberate and attention-demanding than explicit rule use (Palmeri and Gauthier 2004: 300). This is as Dreyfus's model of expertise would predict – like any other type of expert, perceptual experts in a given domain can use their vast repertoire of holistic templates to intuitively recognize novel perceptual stimuli, whereas perceptual novices, like any other type of novice, would have to consciously consult general rules in order to recognize the same stimuli (see Dreyfus and Dreyfus 2004).

The evidently automatic and unconscious character of expert object recognition would give Dreyfus further reason to characterize perceptual expertise as being mindless. For Dreyfus, the automatic and unthinking character of absorbed activity suggests that the motivating forces driving that activity are unthinkable, and hence non-conceptual. Perceptual expertise can viewed as just such an automatic and unthinking activity, being that the processes underlying expert object recognition largely occur outside of the expert's conscious awareness. In the same way as a tennis player's body unconsciously responds to the familiar solicitation of an incoming serve, the expert radiologist's eyes unconsciously responds to the familiar solicitations of an x-ray – the radiologist herself may have no awareness of how her eyes automatically saccade across an image and immediately fixate on a target. Similar to how chess experts who fall pray to the Einstellung effect do not reliably report how their attention is actually being deployed (see Bilalić et al. 2008)[4], expert radiologists' reports of their own visual search methods often diverge from how their eyes are actually scanning an x-ray (Reingold and Sheridan 2011: 534). Experts may also fail to be consciously aware of what they in fact recognize. For example, studies of radiologists have found that the most common form of false-negative error, where an

---

4    The "Einstellung effect" refers to cases where a given problem automatically triggers certain cognitive states and habitual responses that prevent one from detecting a better solution. Bilalic et al. studied the occurrence of this effect among expert chess players, by presenting them with a board situation where checkmate could be reached through a familiar five-move sequence or an uncommon three-move sequence. The experts easily found the five-move sequence; but even when they reported that they were still looking for another solution, they were observed through eye-tracking technology to still be focusing on the features of the problem that were relevant to the solution they had already given. The conclusion drawn was that the problem activated a memory-schema which directs the experts' attention to the features relevant to the familiar sequence, thereby distracting them from the features that would be relevant to another, more simple solution.

abnormality on an image fails to be reported, was one in which the radiologists' eyes fixated on the abnormality for a relatively long duration, suggesting that the abnormality was being recognized – and yet the radiologists consciously decided that no abnormality was present. (Ibid.: 540).

Holistic processing in particular has been shown to be automatically employed by perceptual experts, again as Dreyfus's model would anticipate. Apparently, experts become so attuned to parsing trained stimuli as integrated wholes that they are apparently unable to "turn off" such a holistic processing strategy when trained stimuli are present. Patterns of neural activation in FFA – the region of the visual stream thought to be the locus of holistic processing – can be automatically triggered even when experts are passively viewing trained stimuli without any instruction to perform a recognition task (Richler, Wong, and Gauthier 2011: 131; Tarr and Gauthier 2000). In fact, under certain conditions, the automatic reliance on holistic processing will actually diminish expert performance on certain recognition tasks.[5]

In sum, perceptual expertise seems to instantiate many of the features that Dreyfus attributes to an expert's "perceptual repertoire": Perceptual experts have an increased sensitivity to subtle similarities and differences between perceptual patterns; this sensitivity is directly tied to both low-level and high-level areas of visual/non-cognitive processing; and this processing seems to be holistic, rapid, automatic, unconscious, non-deliberative, and minimally attention-demanding.

**3. The roles of top-down knowledge and contextual associations in expert object recognition**

Does the automatic, unthinking character of perceptual expertise entail that it is exclusively guided by unthinkable, non-conceptual forces? Prima facie, a positive answer seems unwarranted in light of the many cognitive factors that guide especially real-world forms of expert object recognition. Unlike expertise with faces, becoming a perceptual expert in a real-world domain typically involves the long-term acquisition of large amounts of declarative knowledge. This sort of knowledge acquisition is integral to developing perceptual expertise with respect to birds (James and Cree 2010), cars (Gilaie-Dotan et al. 2012), chess (Gobet 2005), wine (Hughson and Boakes 2002), and many other types of objects. Knowledge additionally influences expert visual processing in conjunction with a number of top-down cognitive factors, including task-relevant expectations and goals, semantic memory, and endogenously controlled attention. The presence of these factors together enable experts to make split-second perceptual categorizations, and recognize subtle perceptual cues that novices fail to detect.

For example, along with the exposure to and storage of thousands of radiologic images in long-term visual memory, a radiology student will also acquire a vast amount of straightforwardly conceptual knowledge about anatomy, diseases, and so on. Such knowledge is vital because the perceptual features relevant for correct categorization might not be readily identifiable unless one knows certain relevant information. With the knowledge that a patient is a

---

5    See Gauthier and Bukach 2007 for discussion of "composite task" tests that measure the interference to expert recognition caused by the holistic processing of incongruous images.

gymnast, hurdler, or long jumper, and that subtle pelvic fractures are common injuries for such athletes, a radiologist can better detect these fractures on an x-ray; otherwise, these fractures may pass unnoticed because they are not themselves visually salient (Donovan 2010: 120-1). And even if a stimulus has been detected, knowledge is necessary for properly classifying it: Knowing a patient's case history, facts about anatomy, and even whether the x-ray was underexposed are all important for determining whether a white spot indicates the presence of a lung tumor, a bone, or just a byproduct of the imaging procedure (Wisniewski and Medin 1994: 228).

Another way in which knowledge exerts a top-down influence on perception is through the generation of contextual expectations and predictions. Not only do the features of a certain kind of object appear together in typical, holistic configurations, but objects themselves can appear in typical configurations with other objects. The objects we encounter in everyday experience are seldom perceived in isolation; instead, they are often located in environments in which they bear a semantically coherent relation to other objects – e.g., a microwave is typically seen in a kitchen; a hairdryer is typically seen in a bathroom. With an understanding of these contextual associations, we form expectations about the kinds of objects we may perceive in a given scene, as well as where they may be located, how they may be oriented relative to each other, and so on (Bar 2004: 619). These contextual expectations make the process of object recognition more efficient – for instance, when presented with a familiar scene such as a kitchen, subjects more rapidly recognized a contextually related object like a loaf of bread than an incongruous object like a mailbox or drum (Palmer 1975). Context also helps to resolve perceptual ambiguities: The same amorphous shape may be identified as a car on the street or as someone's shoe, depending on the scene in which the shape is presented, and the expectations one would have about what sorts of objects would be typically found there (Oliva and Torralba 2007; Bar 2004). Overall, contextual expectations assist experts in managing the complexity of the visual environment: Using knowledge of what sorts of objects are typically found in a certain complex scene, the visual system is better able to group and segment elements of that scene into identifiable objects (Gilbert and Li 2013).

The neural regions associated with contextual knowledge and expectations have been found to exert a top-down influence on both the holistic extraction of gist and the rapid recognition of objects. In the same way that the association of facial features enables faces to be processed as holistic units, the contextual associations that experts learn through past experience allow them to form a global impression of a scene and its meaning (Cheung and Bar 2012: 159). In turn, this global impression speeds up object recognition by helping to form a prediction about the potential identity and location of objects in the scene. According to research by Moshe Bar and colleagues (Bar et al. 2006; Kveraga, Ghuman, and Bar 2007), the process of gist perception is initiated by the extraction of low spatial frequency information from a stimulus, i.e., information which does not represent distinctly individuated objects in sharp detail. This information is projected from early visual areas directly and rapidly – at about 130 ms after stimulus onset – to the prefrontal cortex (PFC), and specifically to the orbitofrontal cortex, via the magnocellular pathway of the dorsal visual stream.[6] Notably, only meaningful stimuli, i.e.,

---

6    The orbitofrontal cortex has been associated with a wide range of cognitive functions, the most relevant for our

those stimuli resembling objects associated with category- and identity-relevant semantic memories, were found to activate the orbitofrontal cortex; no activation was found for meaningless visual gratings presented with low spatial frequency (Chaumon et al. 2014).

Signals are then projected from the prefrontal cortex to the inferior temporal cortex (ITC), a high-level area of the ventral stream which contains the fusiform gyrus, and is associated with representing the complex, viewpoint-invariant structures and category identities of perceived objects. Once stored concepts and contextual associations are activated in PFC, they are projected down to ITC so as to provide an initial interpretation of the scene context, as well as predictions about the most likely identities of the objects present therein. These projections reach the ITC around 50 ms before fine-grained, high spatial frequency information arrives from the early visual cortex (Bar 2004; Bar et al. 2006). There are also corresponding projections to ITC from the retrosplenial cortex and parahippocampal cortex (PHC), regions of the medial temporal lobe associated with the long-term storage of memory chunks (Campitelli et al. 2007) and scene-relevant contextual associations (Aminoff, Kveraga, and Bar 2013). Increased activity in PHC has been detected among perceptual experts as compared to novices in several domains, indicating that experts draw upon non-visual associative knowledge about scenes and contexts – e.g., a birder's knowledge of a painted finch's habitat and what its calls sound like, or a radiologist's knowledge of anatomical relations – in order to better recognize objects (see Cheung and Bar 2012).

The effect of these top-down signals is to bias the competition between alternative interpretations of bottom-up visual information, promoting those object-interpretations which are more likely to be accurate given the context, and suppressing unlikely and irrelevant interpretations. These effects are also transferred all the way down to earliest stages of perception: Bottom-up responses in the primary visual cortex (V1) that are incongruent with prior expectations are suppressed, resulting in enhanced or "sharpened" representations with increased information content of expected stimuli (Kok, Jehee, and de Lange 2012). Additionally, prior expectations have been found to evoke a feature-specific pattern of activity in V1 corresponding to the detection of a certain stimulus, even when that stimulus is unexpectedly not present (Kok, Failing, and de Lange 2014). The upshot of these findings is that, by restricting the set of possible interpretations that the visual system has to consider, top-down context predictions lead to more refined and rapid object recognition than what could be achieved in the absence of prior knowledge.

## 4. Dreyfus's feedforward model of perceptual expertise

At this point, Dreyfus could respond by denying any incompatibility between the mindless nature of perceptual expertise and its dependence upon overtly cognitive factors like semantic knowledge and contextual expectations; in fact, he could readily assimilate these cognitive factors into the essentially non-conceptual background understanding that orients all

---

purposes being the processing of affective value and reward, decision-making, guessing and hypothesis-testing, and the formation of expectations. See Bar et al. 2006: 453 for discussion of how the rapid detection of coarse-grained gist and the formation of top-down predictions about object-identities may subserve these cognitive functions in a way that would have conferred evolutionary benefits.

expert coping. According to Dreyfus, our pragmatic engagement with the world relies on our understanding the holistic context of environmental forces that afford us possibilities for action and imbue the world with normative significance. To illustrate this background understanding, he cites Heidegger's example of a professor whose familiarity with the physical and social context of a lecture hall orients his recognition of the blackboard as being out of place. Behind this simple recognition is a background understanding of not only the lecture hall as a physical space affording motor activity, but also its social context, and all the human perspectives and purposes which constitute it: The board is badly positioned relative to the students in the audience who want to see what's being written; it's badly positioned relative to the professor writing on it who wants the students to see what is being written, and so on (Dreyfus 2013: 30). Though this contextual knowledge might seem to be conceptual in nature, it is actually akin to the straightforwardly bodily skills that ground our absorbed interaction with the world – both are forms of non-propositional know-how, and both mindlessly operate in the absence of conceptual thought. Absorbed coping is directly guided by this collective set of background forces, all of which escape apprehension in thought and remain immune to propositional articulation; as Dreyfus writes, "The familiar forces we are absorbed in when we make the judgment that the blackboard is badly placed are not made up of propositional structures to which we can affix bits of language" (Ibid. 20-21). The expert's holistic knowledge of the normative context of action is hence one of the forces that enables the world to stand as a "familiar field of relevant affordances directly soliciting our responses" (Ibid. 17).

Coming now to perceptual expertise, the model of perception that Dreyfus endorses can similarly explain how abstract background knowledge shapes the perceptual expert's responses to environmental affordances without ever rising to the level of conscious, conceptual thought. To explain how the environment can directly solicit an absorbed coper's responses, Dreyfus invokes J.J. Gibson's (1979) ecological theory of direct perception, which holds that the information required for a perceiver to experience and engage with the world is entirely contained within the "ambient optic array," or the structured patterns of light received from the environment by the retina. Our experience of stable three-dimensional objects arises through the detection of perceptual invariants within the optic array, that is, spatiotemporal patterns of stimulation that remain constant while other parts of the array change due to the perceiver's bodily movement. One type of perceptual invariant is an affordance, which refers to the various possibilities for action that objects in the environment offer to a perceiver. Gibson's notion of affordances is thus useful to Dreyfus because it can supplant reasons in explanations of action: Different situations reliably elicit a common pattern of response from agents not because they uniformly provide a set of cognitively appreciable reasons for action, but because they share perceptually available invariant structures to which agents can respond without the mediation of cognitive reasons or other mental representations.

Dreyfus explains the perceptual system's "mindless" interaction with environmental affordances by invoking a model of "feedforward simulated neural networks" (2002: 374-7; 2005: 54-55). The neural network would be comprised by multiple layers of feature detectors, organized hierarchically in increasing degrees of abstraction. Nodes in each layer are responsible for detecting the presence of certain patterns within the input from lower-level nodes. The

highest level of the network could be abstract enough to detect those features in the ambient optic array that indicate the overall significance of a situation. The network's final output would correspond to the response that the situation solicits. Dreyfus's claim is that, through repeated practice and the reinforcement of its correct responses, such a network could learn to reliably discriminate certain stimuli. The fact that the network can learn suggests a way of understanding perception in general as being informed by past experience without that experience being stored as specific memories or associated with conceptual rules. The network needs not store particular memory representations with which current perceptual representations are compared and judged to be similar or dissimilar. Rather, past experience influences present perception through strengthening the connections between neural nodes in the network, such that certain inputs and outputs become more tightly paired together. Through a process of Hebbian learning, where the activity of one node or neuron becomes increasingly synchronized with the activity of another, similar inputs will come to produce the same or similar output.

Whatever way in which a neural network or an embodied human perceiver ultimately comes to detect practically relevant similarities among the invariant features present within an ambient array,[7] Dreyfus's overarching point is these features need not be available to the mind: The nodes in the neural network responsible for directly picking high-order invariants such as affordances remain hidden from the view of the perceiving agent. There may even be nodes tuned by past experience which function as contextual expectations and background knowledge, but they too are hidden from view – all an agent observes is that a certain input solicits a certain response. The agent is unable to consciously represent, name, or think of those invariant features which are detected by the brain in soliciting that response. As Dreyfus summarizes,

> Gibson's account of our direct pick-up of affordances as high order invariants in the optic array, and neural net considerations as to how the brain might detect such invariants, suggest that expertise does not require concepts. Indeed, the basis of expert coping may well be the sort of features that the expert could not be aware of and would not be able to think. (2005: 58)

Being unthinkable and unconscious, these invariant features cannot be brought into a McDowellian "space of reasons," that is, they cannot be taken as reasons for justifying how the perceptual expert categorizes what is seen. Perceptual experts on Dreyfus's view have no conscious access to the abstract, higher-order features that ground their skillful recognition of objects; accordingly, they would be unable to retrospectively reconstruct the reasons why they categorize the objects in the ways that that they do. For Dreyfus, this inability makes perceptual expertise, along with all expert coping, not even an implicitly rational activity.

---

7    Dreyfus (2002: 376) mentions three ways in which, without any conceptual rules, the network could learn to process disparate inputs as being relevantly similar for producing a certain output response: There could be innate, non-cognitive gestalt structures that group inputs together; the temporal order and frequency of inputs could come to signify a shared relevance for a certain output-response, like how nearby objects which afford a reaching response would be detected more early and often than things which do not afford reaching; and inputs could be grouped as similar according to whether they tend to produce a practically satisfactory response.

**5. Against a "mindless" account of perceptual expertise: attention, memory, and conscious agency**

While there is some merit to Dreyfus's characterization of perceptual expertise as a process occurring outside of an agent's awareness, we can still resist the implication that this process is essentially non-conceptual and non-cognitive. Even while granting that a perceptual expert's neural network is subconsciously attuned to detecting abstract features relevant for the rapid categorization of objects, we must acknowledge that perceptual expertise is not exclusively passive, unconscious, ineffable, bottom-up, and feedforward. In particular, Dreyfus's model seems designed to exclude the possibility that expert coping can be consciously guided through agent-directed attention. According to Dreyfus, all that the agent would be consciously aware of is the end-result of the stimulus detection process, namely the response that is issued on the agent's behalf. Nowhere within the feedforward network of feature detectors is there room for any agentive influence over the process of stimulus detection; nor does there seem to be a functional analogue for endogenous selective attention. Dreyfus is averse to acknowledging the conscious and cognitive aspects of expertise, as doing so would not square with his account of absorbed coping as being fundamentally mindless, non-cognitive, and non-representational. Yet, as with most other forms of absorbed coping, perceptual expertise is an active process that can involve conscious and cognitive top-down control.

One striking way in which a perceiver can consciously influence the visual processing of objects is by adopting a certain task-relevant intention. This sort of influence is evident in a study by Assaf Harel and colleagues (Harel, Kravitz, and Baker 2014), where subjects were each given a variety of different perceptual identification tasks to be performed while viewing the same stimulus. The tasks were related to either conceptual characteristics of the object or the physical characteristics of the image. In one trial, subjects would be presented with a picture of a cow, say, and would have to answer whether a cow is a man-made or natural object; in another trial, the same cow would be presented and subjects would be asked whether the image of the cow was tilted clockwise or counter-clockwise. What the study found was that there were different patterns of activation corresponding to each task in the posterior fusiform gyrus (pF) of the ventral temporal cortex, as well as the lateral prefrontal cortex (LPFC). In other words, it was not as though the same object-image generated a consistent, bottom-up pattern of activation in the high-level areas of vision, regardless of the task. Rather, the response of these high-level areas to the same object-image varied across each task. The representations in these areas were so task-dependent that, by changing the task context, the ability to decode which object-image was being perceived just from the corresponding pattern of activation in pF (and LPFC) was significantly reduced.

The implication of this study is that object representations in the visual stream can be modulated by conscious, personal-level states of the observer. Unlike past knowledge or context associations, which may passively influence perceptual experience at a subpersonal level without a perceiver's having some say in the matter, a given task context prompts a perceiver to deliberately adopt a corresponding intention or behavioral goal. These consciously selected intentions and goals in turn shape the patterns of neural activation in the visual system. Dreyfus's account of absorbed coping, however, views the conscious representation of goals on the part of

an agent as anathema to skillful performance. Goals are not consciously represented by the absorbed coper, nor are they unconscious representations which the coper could possibly entertain in conscious thought (Dreyfus 2002: 377-8). Nonetheless, we see in Harel et al.'s study that a perceiver's consciously adopting the goal of correctly responding to a given identification task has a direct effect on how both low-level and high-level stages of the visual system respond to the perceived stimulus.[8]

Dreyfus might object that answering questions about an image falls short of being a form of skillful expertise; consequently, even if a perceiver's conscious intent or adoption of a behavioral goal for answering such questions comes to influence perceptual object processing, those sorts of conscious states would not influence the bottom-up operation of the perceptual expert's neural network and her genuinely mindless perceptual skills. Recall that several empirical studies have claimed that perceptual expertise is ultimately an automatic and stimulus-driven process. Due to long-term training which has tuned the response of neurons in the visual cortex to specific stimuli, perceptual experts can't "turn off" their holistic processing of those stimuli (Tarr and Gauthier 2000; Richler, Wong, and Gauthier 2011).

While such a claim would accord well with Dreyfus's mindless, feedforward model of perceptual expertise, it is undermined by competing research showing that perceptual expertise can be agent-driven rather than stimulus-driven, and that the conscious states of an expert perceiver can activate the skills involved in perceptual expertise. Harel et al. (2011) tested the visual recognition abilities of car experts as compared to novices. Subjects were presented with a rapid series of face, car, and airplane images, and were tasked with detecting whether the same image repeated twice in a row. (Notably, successive car and airplane images were to be judged the same if they both showed the same make and model of car or airplane – e.g., "Honda Civic" – regardless of whether the images differed in color, orientation, or even year of production.) In the first experiment, car experts predictably were more accurate than novices in recognizing identical cars, whereas no significant difference in accuracy was observed for airplane images. Moreover, fMRI scans of the car experts' brains revealed widespread, car-selective activation that was distributed across neural areas within and outside of the visual system. When experts recognized cars, increased activity was observed in the early visual cortex, as well as high-level regions of the ventral stream which are responsive to visual objects, semantic categories, and scene-contexts (e.g., the lateral occipital complex, fusiform gyrus, and parahippocampal cortex). There was also activity found in parietal areas such as the precuneus and intraparietal sulcus, as well as the dorsolateral prefrontal cortex; these regions are together implicated in the fronto-parietal dorsal attention network, which is thought to be responsible for the top-down, voluntary,

---

8    It is worth noting that the patterns of activation in the early visual cortex (V1) were relatively task-independent – that is to say, experimenters could determine which object was being seen from the pattern of activation in the early visual cortex regardless of which task was being performed. Harel, Kravitz, and Baker hence write that "in all visual regions it was still possible to decode object identity across tasks, suggesting that although representations are perturbed, they are not completely changed" (2014: 968). We might say, then, that the top-down influence of behavioral goals or observer intent do not construct an object representation out of whole cloth – they may penetrate the visual processing of visual object representations, but (at least in non-hallucinatory cases) they do not fully replace the bottom-up object information delivered from the early visual cortex.

and goal-oriented allocation of attention (Corbetta and Shulman 2002). The results of the first experiment lend support to the hypothesis that the neural basis of perceptual expertise for cars extends across a wide range of non-visual areas in the brain, rather being restricted solely to face-selective visual areas like the fusiform face area. Additionally, the activity of the fronto-parietal attentional network suggests that top-down attentional allocation was enabling the perceptual engagement of car experts with the objects of their expertise.

Harel et al. (2011) employed a second experiment to test the hypothesis that the neural activity underlying the perceptual expertise of car experts could be controlled in a top-down fashion. Car experts and novices were again presented with a rapid series of car and airplane images, and had to respond when they recognized that the same image was immediately repeated. This time, however, subjects were directed to attend only to car images for one half of the trials, and to airplane images for the other half. Now, if it were true that perceptual expertise is an automatic and stimulus-driven skill, then the same patterns of neural activation which car experts evince in detecting pairs of repeated and identical car images should be triggered by those pairs even when cars were not task-relevant, i.e., during trials in which the experts were told to attend only to airplane images. Yet, researchers found the opposite of what would be predicted under the hypothesis that perceptual expertise is automatic and purely stimulus-driven. In the trials where cars were not task-relevant and hence were not the subject of experts' top-down attentional engagement, experts did not display the sorts of car-selective patterns of neural activity that were observed in the first experiment; in fact, their neural responses to the task-irrelevant cars were nearly identical to that of novices. This result suggests that the widespread neural activity characteristic of perceptual experts – activity which undergirds their enhanced abilities for object recognition – is only found in conjunction with the intentional allocation of attention to objects in their domain of expertise. When perceptual experts aren't actively attending to these objects, their expert abilities remains inactive.

We can now contrast Harel's findings about the role of top-down attention and explicit intention in perceptual expertise with Dreyfus's claim that such personal-level, agent-driven states should impede an expert's skillful performance. This claim should hold true even for the skills of perceptual experts as well – if an expert consciously intends to recognize objects in one's domain of expertise by voluntarily attending to them, then the expert's advantage over a novice perceiver should be degraded. Dreyfus hence seeks to explain the expert's perceptual/non-conceptual repertoire of recognitional abilities in such a way that renders conscious control over these abilities unnecessary, if not impossible. Grounding expert perception on the model of a feedforward neural network allows Dreyfus to show how a perceptual expert could skillfully respond to stimuli without the help of conscious representational states. Since the network is exclusively feedforward, there would be no role for top-down feedback from higher layers of the network to lower layers, or from non-perceptual parts of the brain to the perceptual network itself. Dreyfus does acknowledge that there is a feedback loop between the network's output responses and the environment, which allows the network to passively learn from past experiences in a process of trial-and-error reinforcement. Still, not only does the feedforward model lack any mechanism by which personal-level states could directly modulate the operation of the perceptual network, but the information that the network processes, and the manner in

which it produces skillful responses as a result, cannot be consciously represented to a subject; as Dreyfus writes, "Obviously, the sort of knowledge such a system embodies could not be something one was conscious of and so could not be understood as a conscious or unconscious representation" (2002: 383).

However, the dependence of expert object recognition on the voluntary allocation of selective attention gives us further reason to reject a Dreyfusian account of perceptual expertise. Rather than degrading expertise, Harel's studies have shown that personal-level, agent-driven states like intention and attention actually enable the patterns of neural activity that underlie skillful object recognition. When experts do not actively engage their attention in response to the demands of a specific perceptual task, the patterns of activity exhibited in both low- and high-level visual areas do not differ from those of novices – a finding which would not be predicted if Dreyfus were right that attention should play no role in perceptual expertise, and that expert object recognition is a totally mindless, automatic skill exercised outside of an expert perceiver's control.

Moreover, in failing to find a place for the controlled deployment of selective attention, Dreyfus's feedforward model of perception would further fail to account for another aspect of real-world perceptual expertise, namely the flexibility with which perceptual experts can access domain-specific knowledge in order to categorize objects at varying levels of specificity. A number of studies have suggested that perceptual experts automatically process objects at a subordinate level (Gauthier et al. 2000; Tarr and Gauthier 2000); stating this conclusion in terms of Dreyfus's model, once an expert's neural networks have been passively sensitized to detect more fine-grained categories, the expert can't help but effortlessly recognize and discriminate objects under these categories. It is true that subordinate and sub-subordinate category judgments are much easier for perceptual experts to make – for instance, a novice to intermediate birdwatcher might see a bird and think to classify it as a wren, while an expert might see the same bird and think to classify it as a Carolina wren. But, it is not as though in acquiring expertise for at least real-world object domains, perceptual experts are tuned to automatically make subordinate- rather than basic-level judgments, as Dreyfus's model might have it. Otherwise, if subordinate categories replaced basic categories as the default level of judgment for experts, and their subordinate judgments were now automatic, then an expert birdwatcher would make subordinate judgments more efficiently and rapidly than basic judgments – in other words, it would be easier for an expert birdwatcher to see a bird as being a Carolina wren than as simply being a bird. Accordingly, in a wide-ranging study of birdwatchers by Kathy Johnson and Carolyn Mervis, experts were found to be equally efficient in perceptually identifying objects at a basic, subordinate, or sub-subordinate level, depending on task demands (1997: 264-7).

The equal facility of experts with each of these levels of categorization suggests that they can skillfully respond to perceptual tasks by flexibly drawing upon multiple sources of information, drawn in large part from the vast category-relevant knowledge stored in semantic memory. Different information will be pertinent for different levels of classification – e.g., the features which distinguish a white-crowned sparrow from other sparrows would not be sufficient for distinguishing sparrows in general. Perceptual experts will thus have to access different sorts

of category-relevant information in order to know which distinguishing features they should attentionally select as being most relevant to an intended level of classification. For forms of real-world perceptual expertise, the knowledge of category-relevant information will also include knowledge of more abstract features as well as features from other sense-modalities. A birdwatcher in the field will often classify some bird not only on the basis of available visual cues, which may be rather limited in places like a forest, but also by relying on knowledge of where in a forest the bird is most likely to be found, and what its song sounds like. Through efficiently accessing these perceptual and conceptual sources of information, experts are able to deploy their attention to subtle perceptual features that would otherwise not figure as perceptually salient in the absence of that access (Ibid.: 274).

There are several lessons to be drawn for Dreyfus's feedforward model of perceptual expertise. First, perceptual expertise need not be based upon an automatic, mindless recognition of objects at a fixed level of specificity. Real-world experts are instead capable of flexibly responding to various perceptual tasks that each require objects to be classified at different category-levels. Even if Dreyfus were to hold that such flexible responsiveness should be the mark of expertise rather than be incompatible with it, his model of perception doesn't seem to ultimately allow for such flexibility. That is, if Dreyfus were to ground perceptual expertise on the model of perception that supposedly grounds all other forms of expertise, then his account could not acknowledge the fact that perceptual experts can exert some conscious control over the process of object recognition, and further that the knowledge embodied in their recognitional skills is not wholly inaccessible to conscious awareness.

Second, a purely feedforward model of object recognition would fail to explain the top-down influence of selective attention on visual object processing. Over time, a purely feedforward neural network could become attuned to the subtle perceptual patterns that experts rely upon in making visual classifications; through Hebbian learning, the connections between the nodes that detect domain-specific features would be strengthened, and the connections between irrelevant feature-detecting nodes would become inhibited. Even so, what a purely feedforward model misses is how top-down attention actively places a thumb on the scales of visual processing through a mechanism Robert Goldstone (1998: 588-9) calls "attentional weighting": Selective attention can not only strengthen or amplify the processing of category-relevant features, but can also reconfigure the dimensions along which features are processed as belonging to the same category. In acquiring an ability for perceptual categorization, perceptual experts often learn to ignore sensory features which are otherwise perceptually salient, and focus on more subtle features that are better predictive of category membership. Together with the development of other top-down influences like expectations and semantic memory, learning to preferentially attend in a certain way leads to the re-weighting of neural responses in visual areas to category-relevant and irrelevant features (Gilbert and Li 2013). In turn, attentional weighting contributes to the reshaping of perceptual similarity space, and the sharpening of perceptual category distinctions. By attending to stimuli within the same category (e.g., color), the perceptual features on that dimension will become stretched relative to features on the unattended dimension (e.g., shape), meaning that their differences from the features on the irrelevant, unattended dimension will become sensitized. As a result, selective attention

contributes the development of categorical perception effects whereby intra-category similarities and inter-category differences between stimuli become more perceptually salient (Goldstone and Byrge 2015: 820; see also Nosofsky 1986; Smith and Heise 1992). Ultimately, the power of selective attention and cognitive factors to reweight the neural responses of perceptual systems gives us further reason to think that perceptual expertise cannot be encapsulated within a purely feedforward network, immune to conceptual and conscious influence. As Goldstone and Byrge conclude, "We humans do not simply base our categories on the outputs of perceptual systems independent of feedback. Instead, our perceptual systems become customized to the task-useful categories that we acquire…. [The] fast and widely prevalent recurrent connections from higher to lower cortical regions makes it difficult, sometimes impossible, to identify a 'forward-volley' stage of sensory processing that is uninfluenced by attention" (2015: 821).

In sum, we have shown that concepts, memory, and attention – the three things which Dreyfus claims should not be involved in "mindless" expertise – are actually integral to real-world perceptual expertise. A central part of Dreyfus's non-conceptualism is the view that expert intuition rests on a purely perceptual repertoire of abilities for discriminating a vast array of stimuli and situations; yet, we have seen how the model of perception which is supposed to instantiate these abilities is fundamentally flawed. The neural activity underlying perceptual expertise is widely distributed in the expert perceiver's brain, extending beyond purely perceptual areas and into areas associated with cognition, memory, and top-down selective attention. Moreover, this distributed activation does not simply indicate that perceptual nodes are passing along their outputs to higher, abstract levels in the network – rather, conceptual information from cognitive areas actively shapes the outputs of perceptual areas. In contrast to previous accounts which have suggested in a Dreyfusian vein that expert object recognition is localized in higher-level areas of the visual stream like the fusiform face area, the work of Harel and others offers strong evidence that visual areas are the site at which bottom-up and top-down signals are integrated, and where processes underlying both conceptual and perceptual expertise come to overlap. As Thomas James and George Cree suggest, "If, as we argue, objects are not just processed using visual information, but also conceptual knowledge associated with the object, then perhaps the fusiform gyrus does not represent a purely perceptual stage in visual processing, but instead represents a conceptual stage of object processing" (2010: 348).

Against Dreyfus, then, we may conclude that perceptual expertise ultimately relies a great deal on the expert's conceptual repertoire. This repertoire contains elements which may be uncontroversially recognized as concepts: Advanced perceptual classification requires that experts have learned and stored in semantic memory a vast amount of knowledge concerning their domains of expertise. The conceptual repertoire would also incorporate top-down cognitive factors like expectations, context associations, and task-relevant intentions. Together with trained selective attention, these factors make it possible for experts to skillfully and rapidly recognize objects in ways that outstrip the untutored perception of novices. In the final section, I will briefly consider how this integrative view of perceptual expertise coheres with another model of expert activity that similarly rejects Dreyfus's "mindless" account of intuitive absorbed coping. Applying this model to perceptual expertise, we can understand how the conceptually modulated

operation of attention and memory can turn perception itself into an activity which is both skillful and rationally minded.

## 6. Conclusion: Towards a "Minded" Account of Perceptual Expertise

The necessary involvement of concepts, attention, and memory in object recognition entails that even perceptual expertise cannot be adequately captured by Dreyfus's non-cognitive and non-conceptual account of absorbed coping – this in spite of the fact that expert intuition must ultimately be grounded for Dreyfus upon the types of skills involved in expert object recognition. We may instead look to incorporate perceptual expertise within other theoretical models of expert activity that integrate cognitive factors with the automatic skills that Dreyfus takes to define absorbed intuitive coping.

One such model is offered by Sutton et al. (2011), which understands the relation between cognitive control and automaticity from within a framework of "Applying Intelligence to the Reflexes," or "AIR" (2011). In more recent work that lays out the theoretical space of debates on automaticity and cognition in skilled action, Christensen, Sutton, and McIlwain (2016) develop what they call a "Mesh" theory, proposing that cognitive and attentional control is highly integrated with automatic motor processes. Though much expertise involves mastering skills to the point of being automatic habits, experts often perform in unpredictable contexts with a great number of dynamic variables – in these contexts, totally automatic, inflexible responses will be sub-optimal (see also Toner, Montero, and Moran 2015). So, the AIR model posits that experts can access consciously reconfigure their otherwise unconscious, automated, and stably chunked patterns of behavior into flexible responses to new situational contingencies (Ibid: 96). Exerting cognitive control over automatic behavioral routines need not totally disrupt absorbed coping and revert the expert back to a novice-like stage of self-conscious deliberate activity, as Dreyfus would have it. If this process of cognitive control goes smoothly, then one may indeed feel like "a mindless Dreyfusian expert" who seamlessly remains within the flow of absorbed activity, even though one is actually "mindfully engaging in both paying attention to the demands of a particular performance moment and the most efficient way in which to retrieve chunked material in order to effectively meet these demands" (Geeves et al. 2013: 10).

Similarly, the activity of expert object recognition can be understood as relying on automatic perceptual processes which are nonetheless integrated with cognitive factors under a perceiver's control. On the one hand, an expert's pattern of saccadic eye movements are largely automatic: For instance, expert radiologists are generally unaware of how their eyes immediately saccade across an x-ray and locate an abnormality, even though their pattern of eye movements (e.g., relatively few but long fixations) is partly what distinguishes them from novice radiologists looking at the same image. Additionally, the subpersonal neural mechanisms underlying expert object recognition are essentially inaccessible to conscious awareness. On the other hand, whereas Dreyfus might take the automaticity and conscious inaccessibility of expert eye movements and neural processing to demonstrate the "mindless" character of perceptual expertise, we have seen from Harel's research how consciously adopted, task-relevant intentions

actually enable the patterns of neural activity responsible for the expert's enhanced perceptual abilities.

As for automatic eye movements, these too are linked to the top-down influence of cognitive states and attentional engagement. Experts may not be deliberately controlling each saccade across a visual scene, or even intending to move their eyes in a certain way, but their eye movement patterns are still part of their intentional response to task demands. In a classic study by Yarbus (1967), a subject was presented with Repin's painting "The Unexpected Visitor," and was given tasks like remembering what clothes were worn by the people in the painting, or estimating how long the visitor entering the room had been away. The resulting eye movement patterns make sense as rational responses to the semantic content of these tasks: when asked to remember the clothes, the subject's eyes moved up and down the individual bodies, whereas when estimating how long the visitor was away, the subject's eyes swept back and forth from the entering visitor's face to the faces of the people seated in the room, apparently searching their expressions for emotional cues (Fridland 2017: 4343). Again, these saccade patterns were automatic and non-deliberate, but they nevertheless followed from the top-down selective attention was intentionally applied by the perceiver in response to a given task. These patterns are also obviously linked to the perceiver's background conceptual knowledge, whether about clothes or social interactions. This same sort of intelligent responsiveness is observable in the eye movement patterns of experts across a wide array of rarefied and mundane forms of skill, from athletes (Mann et al. 2007), chess players and radiologists (Reingold and Sheridan 2011), to experts at everyday activities like making tea or sandwiches (Land 2006). Dreyfus could potentially object that his feedforward model of perceptual expertise can accommodate top-down intentions – the expert's intention to perform a given task sets in motion automatic eye movements and perceptual processes that, without subsequent cognitive interference or attentional engagement, ultimately issue an appropriate response from the bottom-up. However, it is not as though the expert's top-down instigation of a saccade pattern is a one-off event after which cognitive control ceases – when performing extended, multi-stage tasks, expert copers continue to be attentively engaged in monitoring their activity, and as a result are constantly altering their saccade patterns to serve different functions – e.g., locating an object, guiding action with respect to that object, and checking that the action was successful – as they move from one stage to the next (Land and Hayhoe 2001: 3563).

In fact, eye movements are a type of motor activity like any other, and so they can also be incorporated within an "AIR" model of perceptual expertise. As with other skilled motor routines, the automation of saccade patterns frees up perceivers to strategically deploy their attentional resources for the sake of achieving higher-order objectives (e.g., "identify the species of wren," "diagnose the cancer"). It is precisely because experts don't have to directly attend to their saccade routines that they can instead attend to the subtle perceptual features required for task-relevant object recognition. At the same time, maintaining top-down control over perceptual and attentional routines allows perceptual experts to flexibly access multiple forms of sensory information present in a perceived scene, as well as chunked perceptual and semantic information stored in memory (Harel 2016: 91). Extending the idea of "applying intelligence to the reflexes" involved in object recognition, we hence might say that automatic perceptual

reflexes are intelligible in light of the conceptual knowledge and selective attention that experts intelligently apply in responding to a perceptual task.

For his part, Dreyfus crucially recognized that perceptual expertise serves as the basis for most every form of absorbed coping. But, he was wary of allowing cognitive factors like attention and memory to figure in his account of expert activity, because he viewed them as being the first steps toward ejecting experts from the flow of absorbed coping, and transforming their mindless, non-conceptual engagement with the world into an object of detached conceptualization and discursive reflection. However, the integrative account of perceptual expertise defended here suggests that concepts, attention, and memory are present even in the perceptual foundations of seemingly "mindless" coping and intuition. If we incorporate expert perceptual coping within more plausible models of expertise such as "AIR," then we can better understand how skillful object recognition is on a par with other expert skills, insofar as their optimal functioning is compatible with, and to some degree enabled by, the top-down influence of conceptual knowledge and attentional engagement. Indeed, these cognitive factors ultimately allow an expert's skillful activity to be flexibly and intelligently responsive to the world and its normative demands. We may thus conclude that cognitive and attentive awareness within the flow of expert perceptual coping would make the expert's experience of the world both mindful and rationally minded.

## Acknowledgments

## References

Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences*, *17*(8), 379–390.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–454.

Bilalic, M., McLeod, P. & Gobet, F. 2008. Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*(3), 652–661.

Bilalić, M. (2016). Revisiting the role of the fusiform face area in expertise. *Journal of Cognitive Neuroscience*, *28*(9), 1345–1357.

Bläsing, B. E., Güldenpenning, I., Koester, D., & Schack, T. (2014). Expertise affects representation structure and categorical activation of grasp postures in climbing. *Frontiers in Psychology, 5*(1008), 1-11.

Breivik, G. (2013). Zombie-like or superconscious? A phenomenological and conceptual

analysis of consciousness in elite sport. *Journal of the Philosophy of Sport*, *40*(1), 85–106.

Campitelli, G., Gobet, F., Head, K., Buckley, M., & Parker, A. (2007). Brain localisation of memory chunks in chess players. *International Journal of Neuroscience*, *117*(2), 1641–1659.

Chaumon, M., Kveraga, K., Barrett, L. F., & Bar, M. (2014). Visual predictions in the orbitofrontal cortex rely on associative content. *Cerebral Cortex*, *24*(11), 2899–2907.

Cheung, O.S., and Bar, M. (2012). Visual prediction and perceptual expertise. *International Journal of Psychophysiology, 83*(2), 156–163.

Christensen, W., Sutton, J., & McIlwain, D. J. F. (2015). Putting pressure on theories of choking: towards an expanded perspective on breakdown in skilled performance. *Phenomenology and the Cognitive Sciences*, *14*(2), 253–293.

Christensen, W., Sutton, J., & McIlwain, D. J. F. (2016). Cognition in skilled action: Meshed control and the varieties of skill experience. *Mind & Language*, *31*(1), 37–66.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215.

Donovan, T. (2010). Attention and medical diagnosis. In E. B. Goldstein (Ed.), *Encyclopedia of Perception* (pp. 119–121). Los Angeles: SAGE.

Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. New York: Harper & Row.

Dreyfus, H. L. (2002). Intelligence without representation: Merleau-Ponty's critique of mental representation. *Phenomenology and the Cognitive Sciences*, *1*(4), 367–383.

Dreyfus, H. L. (2005). Overcoming the myth of the mental: How philosophers can profit from the phenomenology of everyday expertise. *Proceedings and Addresses of the American Philosophical Association*, *79*(2), 47–65.

Dreyfus, H. L. (2007a). The return of the myth of the mental. *Inquiry*, *50*(4), 352–365.

Dreyfus, H. L. (2007b). Response to McDowell. *Inquiry*, *50*(4), 371–377.

Dreyfus, H. L. (2013). The myth of the pervasiveness of the mental. In J.K. Schear (Ed.), *Mind, reason, and being-in-the-world: The McDowell-Dreyfus debate* (pp. 15–40). London: Routledge.

Dreyfus, H. L., & Dreyfus, S.E. (1988). *Mind over machine: The power of human intuition and expertise in the era of the computer* (2nd ed.). New York: Free Press.

Dreyfus, S. E. (2004). The five-stage model of adult skill acquisition. *Bulletin of Science, Technology & Society, 24*(3), 177-181.

Fahle, Manfred. 2002. Introduction. In M. Fahle & T. Poggio, *Perceptual learning* (pp. Ix-xx). Cambridge: MIT Press.

Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 41–67). New York: Cambridge University Press.

Fridland, E. (2017). Automatically minded. *Synthese*, *194*(11), 4337–4363.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103*(2)*,* 322–330.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–197.

Geeves, A., McIlwain, D. J. F., Sutton, J., & Christensen, W. (2014). To think or not to think: the apparent paradox of expert skill in music performance. *Educational Philosophy and Theory, 46*(6), 674–691.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilaie-Dotan, S., Harel, A., Bentin, S., Kanai, R., & Rees, G. (2012). Neuroanatomical correlates of visual car expertise. *Neuroimage, 62-334*(1), 147–153.

Gilbert, C. D., and Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience, 14*(5): 350–363.

Gobet, F. (2005). Chunking models of expertise: Implications for education." *Applied Cognitive Psychology, 19*(2), 183–204.

Gobet, F. & Chassy, P. (2009). Expertise and intuition: A tale of three theories. *Minds and Machines 19*(2), 151–180.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*(1), 585–612.

Goldstone, R. L., & Byrge, L. (2015). Perceptual learning. In M. Matthen (Ed.), *The Oxford handbook of philosophy of perception* (pp. 812-832). Oxford: Oxford University Press.

Grill-Spector, K. Object perception: Physiology. In E. B. Goldstein (Ed.), *Encyclopedia of Perception* (pp. 648–653). Los Angeles: SAGE.

Harel, A. (2016). What is special about expertise? Visual expertise reveals the interactive nature of real-world object recognition." *Neuropsychologia, 83*, 88–99.

Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences, 111*(10), 962–971.

Harel, A., Ullman, S., Harari, D., & Bentin, S. (2011). Basic-level categorization of intermediate complexity fragments reveals top-down effects of expertise in visual perception. *Journal of Vision, 11*(8), 1–13.

Hughson, A. L., & Boakes, R. A. (2002). The knowing nose: The role of knowledge in wine expertise. *Food Quality and Preference, 13*(7–8), 463–472.

James, T. W., & Cree, G. (2010). Perceptual and conceptual interactions in object recognition and expertise. In I. Gauthier, M. J. Tarr, & D. Bub (Eds.), *Perceptual expertise: Bridging brain and behavior* (pp. 333-352). New York: Oxford University Press.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126*(3), 248–277.

Kellman, P. J., & Garrigan, P. (2009). Perceptual learning and human expertise. *Physics of Life Reviews, 6*(2), 53–84.

Karni, A., & Sagi, D. (1995). A memory system in the adult visual cortex. In B. Julesz & I. Kovacs (Eds.), Maturational windows and adult cortical plasticity (pp. 95–111). Reading: Addison-Wesley.

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience, 26*(7), 1546–1554.

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron, 75*(2), 265–270.

Kveraga, K., Ghuman, A., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition, 65*(2), 145–168.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25*(3), 296–324.

Land, M. F., and Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25-26), 3559–65.

Lu, Z. L., Hua, T., Huang, C. B., Zhou, Y., & Dosher, B. A. (2011). Visual perceptual learning. *Neurobiology of Learning and Memory, 95*(2), 145–151.

Mann, D. T. Y., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport & Exercise Psychology, 29*(4), 457–78.

Montero, B. (2016). *Thought in action: Expertise and the conscious mind*. Oxford: Oxford University Press.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39–57.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520–527.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*(5), 519–526.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Reviews Neuroscience, 5*(4), 291–303.

Raftopoulos, A. (2001). Perceptual learning meets philosophy: cognitive penetrability of perception and its philosophical implications. In J. Moore & K. Stemming (Eds.), *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 802–808). Mahwah: Lawrence Erlbaum.

Reingold, E. M., & Sheridan, H. (2011). Eye Movements and Visual Expertise in Chess and Medicine. In S. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook on eye movements* (pp. 528–550). Oxford: Oxford University Press.

Richler, J. J., Wong, Y. K., & Gauthier, I. (2011). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science, 20*(2), 129–134.

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In B. Burns (Ed.), *Percepts, concepts and categories: The representation and processing of information* (pp. 233–272). Amsterdam: Elsevier.

Sutton, J., McIlwain, D., Christensen, W., & Geeves, A. (2011). Applying intelligence to the reflexes: Embodied skills and habits between Dreyfus and Descartes. *Journal of the British Society for Phenomenology, 42*(1), 78–103.

Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3*(8), 764–769.

Toner, J., Montero, B. G., & Moran, A. (2015). The perils of automaticity. *Review of General Psychology, 19*(4), 431–442.

Toner, J., Montero, B. G., & Moran, A. (2016). Reflective and prereflective bodily awareness in skilled action. *Psychology of Consciousness: Theory, Research, and Practice, 3*(4), 303–315.

Wang, R., Wang, J., Zhang, J. Y., Xie, X. Y., Yang, Y. X., Luo, S. H., Yu, C., & Li, W. (2016). Perceptual Learning at a Conceptual Level. *The Journal of Neuroscience*, *36*(7), 2238-

2246.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science, 18*(2), 221–281.

Wulf, G. (2013). Attentional focus and motor learning: a review of 15 years. *International Review of Sport and Exercise Psychology, 6*(1), 77–104.

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.