



Seeing the forest for the trees

Scene perception and the admissible contents of perceptual experience

Tom McClelland^a  (tom.mcclelland57@gmail.com)

Abstract

Debates surrounding the high-level contents of perceptual experience focus on whether we perceive the high-level properties of visual objects, such as the property of being a pine tree. This paper considers instead whether we perceive the high-level properties of visual scenes, such as the property of being a forest. Liberals about the contents of perceptual experience have offered a variety of phenomenal contrast cases designed to reveal how the high-level properties of objects figure in our visual experience. I offer a series of equivalent phenomenal contrast cases intended to show how the high-level properties of visual scenes also figure in visual experience. This first-person evidence of high-level scene perception is combined with third-person evidence from the extensive empirical literature on scene categorisation. Critics of liberalism have attempted to deflate existing phenomenal contrast cases by explaining the contrasts in terms of non-perceptual contents or in terms of attentional changes. I show that neither response is applicable to my contrast cases and conclude that we do indeed perceptually experience the high-level properties of visual scenes.

Keywords

Attention · Gist · Mental representation · Perception · Vision

1 The admissible contents of perceptual experience

Perceptual experiences have content – a perceptual experience can be accurate to the extent that the world is as the experience represents it to be (Pautz, 2009). A key challenge for our understanding of perceptual experience is to determine its *admissible contents* (Macpherson, 2011). If we consider visual experience, we find there are some properties that we can see – such as redness and roundness – and others that we cannot – such as market value and radioactive decay (Butterfill, 2009). The challenge is to work out which properties can figure in visual experience and

^aUniversity of Cambridge

which cannot. There are a set of ‘low-level’ properties that are generally agreed to be represented in visual experience, including colours, textures, shapes and locations.¹ Conservatives say that the content of visual experience is *exhausted* by such properties, e.g. Briscoe (2015), Brogaard (2018), Carruthers & Veillet (2011), Price (2009), Prinz (2012). Liberals say that besides these low-level properties, visual experience also represents various ‘high-level’ properties, e.g. Bayne (2009), Bayne (2016), Fish (2013), Macpherson (2011), Masrour (2011), McClelland (2016), McClelland (2019), Siegel (2006), Siegel (2012). The list of candidate properties here is much longer but prominent examples include kinds, causal properties, affordances, aesthetic properties, moral properties and the identity of individuals.

What is the content of my perceptual experience as I look out at the cat in my garden? Conservatives say that my visual experience represents only the low-level properties of the cat, such as its shape and colour. Liberals, on the other hand, are open to my visual experience also representing various high-level properties: kind properties like *being a cat*; causal properties like *making the grass squish*; affordances like *being strokeable*; aesthetic properties like *being pretty*; moral properties like *being a bit evil*; and an identity such as *being Atticus*?² Although the existing literature focuses on whether we perceptually experience such high-level properties of visual objects, I want to shift our attention instead to the high-level properties of visual scenes.³ So rather than probing the visual experience of the cat I want to probe the visual experience of the environment in which the cat is embedded. And to achieve this I want to use both the first-person examination of our visual phenomenology and the rich empirical literature on scene perception.

Scenes are quite different to objects. The garden is not just another item on the list of objects I see. Instead, it is more like a container in which individual visual objects appear. Some contents of my visual experience attribute properties to individual objects, but other contents attribute properties to the environment in which those objects are embedded. Among these contents will be various low-level properties of the scene: I visually experience the garden as having certain colours, shapes and textures. But does my visual experience also encompass high-level properties of this scene? The conservative answer is that no such properties are admissible. The liberal answer is that such properties do indeed figure in the contents of my perceptual experience.

What are the candidate high-level properties of visual scenes? My focus will be on scene categories i.e. the kind properties of scenes. In the case described, I might experience the environment *as outdoors*, *as natural*, or *as a garden*. Other

¹Other properties on the list are spatial relations, sizes, motion, volume, orientation and illumination. Although there are no clear-cut criteria for a property being low-level, we will see that classifying properties is not too hard.

²This example is inspired by Helton’s (2016, p. 851) case of the cat in a hammock.

³Fish (2013, pp. 50–52) mentions high-level properties of scenes briefly. Bayne (2016) explores them in more detail in his wider discussion of gist perception, though I explain where I diverge from Bayne in Section 2.

candidate kind properties of scenes include *being a forest*, *being a beach*, *being a kitchen* or *being a carpark*. These are all categories of environment that cannot be straightforwardly reduced to the properties of objects within that environment. I take these properties to be on a par with the object kinds commonly discussed in the literature such as *being a stethoscope*, *being a tomato* and *being a pine tree*. Besides scene kind properties, there are various other high-level scene properties that might figure in my visual experience. I might perceptually experience the garden as having affordances like *being walk-round-able*, aesthetic properties like *being beautiful*, or a specific identity such as *being my garden*. Causal and moral properties of scenes are a little harder to make sense of, but perhaps we can see the garden as *causing the cat to relax* or even see it as being a *good space* in some quasi-moral sense. But why does it matter whether such high-level scene properties figure in visual experience? The admissible contents debate has a number of background motivations, and each of which provides an incentive to explore the high-level properties of visual scenes.

The first motivation is *phenomenological*. We want to be able to describe the phenomenal character of experience. Determining whether or not high-level properties figure in our perceptual experience is crucial to this descriptive project. Conservatives paint a picture of visual phenomenology with a palette that is strictly limited to low-level properties while liberals paint a picture with a considerably richer palette. If we establish that *catness* figures in our perceptual experience we will have learned something important about what our experiences are like. This motivation also encompasses high-level properties of scenes. Establishing that *garden-ness* figures in our perceptual experience would tell us something interesting about our phenomenology.

Unlike high-level object properties, scene properties could be of particular relevance to the phenomenology of peripheral visual experience. Our perceptual representation of objects outside focal attention is surprisingly sparse (Mack & Rock, 1998). As I focus on the cat, my visual experience tells me a great deal about the cat but specifies very little about the individual trees, plants and tools in the surrounding garden. Nevertheless, my perceptual experience doesn't just present me with the few objects that fall under the spotlight of my attention. Instead, it gives me the impression of a rich visual scene. If our perceptual experience includes high-level scene properties we might be able to explain this impression of richness: although our visual experience of peripheral objects is sparse, our visual experience of the overall scenes in which they are embedded is rich.

The second motivation is *epistemological* (Helton, 2016, p. 853). Having particular perceptual experiences justifies us in having particular beliefs. Determining whether high-level properties of objects figure in those perceptual experiences will help us understand what beliefs they justify and how they do so. How, for instance, is my judgement that the object before me is a cat justified? Conservatives must say that it is justified by my perceiving particular low-level properties of the object. But liberals can say that it is justified more directly by my seeing the ob-

ject *as a cat*. Adding such high-level contents to a perceptual experience might give it an *epistemic upgrade*, providing it with greater power to warrant specific judgements. Interestingly though, such high-level content can also *narrow* the justificatory power of a perceptual experience (Siegel, 2006). A perception of just the low-level properties of an animal in the distance might lend equal warrant to the belief that it is a cat and the belief that it is a badger. But once you perceive it *as a cat* your perceptual experience has made a fine-grained distinction and no longer lends warrant to the badger-belief.

If high-level scene properties figure in perceptual experience this too would have important epistemological implications. Most directly, it might account for how we're warranted in making certain claims about our environment, such as judging *that this is a garden*. Here one's perceptual experience might be epistemically upgraded by its high-level contents. Again though, these more fine-grained discriminations come hand-in-hand with a *narrowing* of the beliefs that one's perceptual experience warrants. More indirectly, it might help account for some of our judgements about objects. Consider a case where I judge an object to be a rake despite having no perceptual capacity to discriminate rakes from non-rakes. Maybe my perception of the scene *as a garden*, combined with my knowledge that rakes are found in gardens, helps warrant my judgement. Context is a crucial factor in justification and scene perception might helpfully be construed as a kind of context perception. This could make a contribution to our epistemic situation quite different to that of high-level object perception.

The third motivation concerns *action*. Our actions are guided by what we perceive. When I stroke the cat, for instance, my action is guided by my perception. Liberals can say that my stroking the cat is based on my seeing it *as affording stroking* (McClelland, 2019). Conservatives, on the other hand, will have to posit more intermediate cognition before I'm in a position to act: I might, for instance, see the low-level properties of the cat then have to *judge* that it is strokeable before acting. In this example the affordance belongs to a specific object – the cat – but not all affordances are attached to individuals (McClelland, 2020). A garden might afford walking, a forest might afford sheltering, a field might afford hunting, a kitchen might afford cooking and a carpark might afford parking. These affordances belong to our environment and our perception of such affordances could play an important role in explaining how we act in such environments. This opens up an aspect of perception-for-action that we don't get from perception of object affordances.

To determine whether we do indeed perceptually experience the high-level properties of scenes, we first need to pin down what exactly it means for a property to figure in the contents of perceptual experience. For a property *F* to be part of the content of a mental representation is for *F* to figure in the accuracy conditions of that representation. But what makes a mental representation of *F* a perceptual experience of *F* rather than some other kind of mental representation of *F*? Here we can introduce two individually necessary and jointly sufficient conditions:

I. The Perceptual Condition: A property *F* is represented in a subject's perceptual experience only if the subject's representation of that property is *perceptual* rather than non-perceptual.

II. The Phenomenal Condition: A property *F* is represented in a subject's perceptual experience only if the subject's representation of that property is *phenomenal* rather than non-phenomenal.

These conditions are independent. Mental representations can be perceptual without being phenomenal e.g. one might unconsciously perceive an advert without experiencing it. And a mental representation can be phenomenal without being perceptual e.g. conscious thoughts or emotions are not perceptions. To determine whether the conditions are met we need to have a grip on both the distinction between perceptual/non-perceptual representation and the distinction between phenomenal/non-phenomenal representation.

Your visual representation of the environment around you is a paradigm case of perceptual representation. A perceptual representation is the product of a process driven by sensory inputs. This perceptual process is typically rapid, automatic, cognitively impenetrable and subject to adaptation effects. Your belief that Paris is more northerly than Madrid is a paradigm case of non-perceptual representation. This representation is likely the product of a slower, voluntary process of reasoning that is responsive to your other beliefs and that is neither driven by sensory stimuli nor subject to adaptation effects. Although this serves to highlight the distinction between perceptual and non-perceptual states, what it does not do is give us necessary and sufficient conditions. The prospects of finding such conditions are poor: for any putative hallmark of perception there are perceptual states that plausibly lack this feature and non-perceptual states that plausibly possess it.

Your conscious awareness of reading this paper would be a paradigm case of phenomenal representation. Phenomenal states are those that there is something it's like to be in for the subject of that state, and non-phenomenal states are those that there is nothing it's like to be in for the subject. Again, there are no uncontroversial necessary and sufficient conditions for mental content being phenomenal. Following Bayne (2016), I'll say that the phenomenal content of a state is content that is suitably *reflected* in its phenomenology. This deliberately broad characterisation is neutral on a number of disputed questions about the relationship between an experience's phenomenology and its content such as whether content should be understood in an internalist or externalist way, whether two states with the same phenomenology can differ in content and whether two phenomenal states with the same content can differ in their phenomenology.

Although we can illuminate the perceptual/non-perceptual and phenomenal/non-phenomenal distinctions, what we cannot do is provide a conceptual litmus test for either. Any proposed test would be too contentious to carry any dialectical weight. There is simply too much disagreement about what it takes for a state to be perceptual and what it takes for it to be phenomenal. But the absence of decisive criteria should not deter us. The important thing is that these distinctions mark

real differences in kinds of mental representation, albeit differences that are hard to define. And we can still find good evidence for some mental representation being perceptual and/or phenomenal, even if that evidence is inevitably defeasible.

We can now frame our target question more precisely: are there mental representations of high-level scene properties that satisfy both the Perceptual Condition and the Phenomenal Condition? In Section 2 I will use empirical findings to argue that there are representations of high-level scene properties that satisfy the Perceptual Condition. In Section 3 I will use phenomenal contrast cases to argue that such representations can also satisfy the Phenomenal Condition. In Sections 4 and 5 I consider and reject some conservative alternatives to this conclusion.

2 Do we perceive high-level properties of visual scenes?

Since the 1970s there has been a great deal of research on ‘gist perception’ – our perception of the overall meaning of the visual scene before us. The most striking feature of gist perception is the incredible speed at which it is achieved. This is captured in the following passage from Oliva:

Remarkably, we are able to interpret the meaning of multifaceted and complex scene images [...] in a fraction of a second [...]! This is about the same time it takes a person to identify that a single object is a face, a dog, or a car [...]. An unmistakable demonstration of the brain’s prowess in visual scene understanding can be experienced at the movies: With a few rapid scene cuts from a movie to form a trailer, it seems as if we have perceived and understood much more of the story in a few seconds than could be described later in the same amount of time. (Oliva, 2013, p. 725)

One of the processes that contributes to our rapid understanding of a scene is how quickly we categorise the *kind* of environment we are seeing. For instance, understanding that the visual scene before us is *a game of beach volleyball* is facilitated by categorising the environment *as a beach*. This has led to a significant body of research specifically concerned with scene categorisation i.e. with our recognition of kinds of environment. Our perceptual categorisation of scenes is best understood as one aspect of gist perception: our perception of gist involves both scene-based gist perception – such as seeing a scene as a forest – and object-based gist perception – such as seeing an object in the forest as an animal (Bayne, 2016, p. 109).

Crucially, the evidence suggests that we do not arrive at a categorisation of a scene by building up from our categorisation of objects. So rather than moving from categorising objects as trees to categorising the scene as a forest, we process scene categories through a distinct stream of visual processing. The neural

‘scene network’ encompasses the parahippocampal place area (PPA), the retrosplenial complex/medial place area (RSC/MPA) and the occipital place area (OPA). Although the exact functional role of these brain regions is not yet fully understood, we do know that these areas respond selectively to scenes (Epstein & Baker, 2019). And this neural network interacts in various ways with the distinct neural processes responsible for object-understanding.

Vision science takes it as given that when we perceive a scene we are perceiving the high-level properties of that scene. The ‘scene understanding’ that they target involves rich semantic categories and there is no question of whether our perception of scenes might be limited just to its low-level properties. Indeed, studies typically define scene perception in terms of the perception of scene kinds. This means the empirical literature never really *argues* for the conclusion that high-level properties of visual scenes can indeed be perceived. How, then, might their findings be marshalled to provide such an argument? I suggest that the speed of scene categorisation strongly suggests that we do indeed perceptually represent scene kinds. Whether we perceptually *experience* scene kinds is a further question confronted in Section 3.

Speed is one of the key criteria used to distinguish perceptual from non-perceptual processes. As mentioned above, there is a wealth of empirical data revealing that scene categorisation is achieved at incredible speed. Greene & Oliva (2009), for example, identified the exposure time required for subjects to achieve 75% accuracy in a scene categorisation task. For each of the scene categories investigated, the required exposure time was well below 100ms. The longest exposure time required was for *rivers* at 67ms and the shortest was for *forests* at 30ms. These trials exploited a masking paradigm to rule out the possibility of images being retained in working memory then processed after the image disappeared. Other studies have used different methods to probe processing speed and have similarly demonstrated that scene categorisation occurs very rapidly (Rousselet et al., 2005).

Does this show that scene categorisation is perceptual? One might apply the criterion that any mental process that is sufficiently rapid is perceptual, while all slower mental processes are non-perceptual. If this criterion were used then scene categorisation would come out as perceptual, but this criterion is hard to defend. Many non-perceptual processes – such as ‘System 1’ cognitive processes – are very rapid (Kahneman, 2011). This means that scene categorisation could be construed as a rapid non-perceptual process. Against such a view, one might argue that scene categorisation is faster than the fastest non-perceptual processes so must still be regarded as perceptual. However, this response would require the principled identification of a speed limit on non-perceptual processes but such a limit is extremely hard to specify in a non-arbitrary way. Speed cannot be regarded as a simple litmus test for perception.

A better strategy is to bracket the question of how quickly a non-perceptual process can occur and instead consider how quickly a non-perceptual process can



Figure 1: Reproduced with permission from Greene & Oliva (2010).

begin. Any non-perceptual categorisation of scenes would need to be driven by input from a perceptual system. As Fish points out “interpretative processes would need something to interpret” (2013, p. 50). For example, rapid judgment that a scene is a forest would need to occur *after* perception of the low-level properties on which this judgement is based. However, the evidence suggests that scene categorisation is achieved *before* many of the low-level details about a scene are perceptually processed (Greene and Oliva, 2009). At 30ms the scene has been categorised as a forest but has not yet even been parsed into discrete objects. As Navon (1977) puts it, we see the forest *before* seeing the trees. This makes it implausible that a non-perceptual process could have received sufficient perceptual input to achieve scene categorisation on this timescale.

I argue that the speed of scene categorisation constitutes good – though defeasible – evidence that scene categories are indeed represented perceptually. Does the empirical literature provide other routes to the same conclusion? It is tempting to conclude that adaptation studies achieve this but I argue that they do not provide what is needed.⁴ In an adaptation study by Greene & Oliva (2010), scene images were organised along various continua such as in figure 1.

Subjects were exposed to images at one end of the continuum, then presented with an ambiguous image from the centre of the continuum and asked to categorise it.⁵ The categorisations made by subjects displayed aftereffects. For instance, after prolonged exposure to an image of a field, subjects displayed bias toward categorising ambiguous images (such as the central image on the continuum above) *as forests*. Conversely, after prolonged exposure to an image of a forest, subjects displayed bias toward categorising the ambiguous images *as fields*.

Although striking, this study does not indicate adaptation to high-level scene categories. Greene & Oliva placed the images above along a continuum of ‘openness.’ The field on the left is at the ‘open’ end of the continuum and the forest on the right is at the ‘closed’ end of the continuum. But openness is a *low-level* property of visual scenes defined as the degree to which bounded regions obscure the visibility of the horizon (Oliva & Torralba, 2006). The experiment’s results are

⁴Studies of the automaticity of scene categorisation are also *prima facie* relevant here (Biederman et al., 1983). However, automaticity is a particularly slippery way of distinguishing perceptual and non-perceptual representation, and it is not clear that the studies reveal the right kind of automaticity to be useful.

⁵The continua themselves were based on the results of previous studies (Oliva & Torralba, 2006).

naturally interpreted in terms of subjects first adapting to the openness of a visual scene and then using openness to guide their categorisation of scenes. When adaptation makes the ambiguous image seem more open, for example, subjects are more likely to categorise it as a field. As such, the study does not provide what is needed. Neurological studies have also revealed adaptation effects in scene perception but a recent review of these studies by Epstein & Baker (2019) concluded that these effects are too ambiguous to reveal the contents of the underlying neural representations.

Empirical findings regarding the speed of scene categorisation strongly suggest that there are representations of high-level scene properties that satisfy the Perceptual Condition. Could the empirical evidence also be used to show that the Phenomenal Condition is satisfied? Evidence from vision science is unlikely to justify such a conclusion. It seems that conservatives will always have space to say that a perceptual representation revealed by a study is non-phenomenal. Consider Briscoe's response to Fish's (2013) empirical case for including *being an animal* as a content of visual perceptual experience. Briscoe argues that an object "may be categorized by the visual system as an animal [...] even though the object is absent from the sensory, presentational component of visual experience" (2015, p. 183). A similar response is available to Bayne's (2016) empirically motivated argument for admitting 'scene gist' into the contents of visual experience. The empirical data cited by Bayne may well show that we perceive such scene properties, but this is not enough to show that such properties are perceived phenomenally. To discern whether our perception of scene categorisation is indeed phenomenal we should turn away from such third-person evidence and consider our perceptual experiences from the first-person perspective.

3 Do we perceptually experience high-level properties of visual scenes?

To determine whether our perception of scene categories figures in our perceptual phenomenology we should use the method of phenomenal contrast. This method requires us to identify experience pairs that are alike with respect to relevant low-level contents but that nonetheless differ in their phenomenology. This phenomenal difference then stands in need of explanation. If the best explanation is that the two experiences differ with respect to their high-level perceptual content, this would constitute good evidence that we perceptually experience high-level properties (Siegel, 2012, p. 88). I divide contrast cases into three groups: revelation cases, skilled recognition cases and aspect-switching cases. I will offer cases in each of these groups that can plausibly be explained in terms of our perceptually experiencing high-level properties of visual scenes. To establish that my explanation of the contrasts is the *best* explanation, I will need to confront some alternative explanations. This challenge will be taken up in Sections 4 and 5.



Figure 2: Reproduced with permission from Teufel et al. (2015). The unmanipulated image can be found in the Appendix.

3.1 Revelation cases

Revelation cases are cases in which we perceptually experience an unclear stimulus and our phenomenology changes when we come to recognise what that stimulus is. Consider figure 2.

At first, many subjects experience this image as a meaningless jumble of black and white patches. Look closely, however, and you'll see that this is actually an image of a baby (the baby's head is in the top-left of the image looking down at something). When you recognise the baby your experience changes – there is a phenomenal contrast between the experience you have at t_1 before you have noticed the baby and the experience you have at t_2 once you have noticed it.

This contrast cannot be readily explained in terms of changes among the low-level properties one visually experiences: it seems that the black and white colour qualities, shapes and textures that we experience remain constant through t_1 and t_2 . This motivates liberals to explain the contrast in terms of one's experience at t_2 representing the high-level property of *being a baby*. Figure 3 shows two scene-based examples that mirror the baby case.

As with the baby case, you might start by experiencing figure 3A as a meaningless array of black and white patches. But when you recognise it *as a waterfall* your experience changes. Similarly, your experience of figure 3B might change when you recognise it *as a cavern*. As with the baby case, there is no obvious change in the low-level contents of your perceptual experience. I suggest that each contrast should be explained in terms of one coming to perceptually experience the image as being of a particular scene kind. On this account of the contrasts, we perceptually experience the high-level scene properties of *being a waterfall* and *being a cavern* respectively.

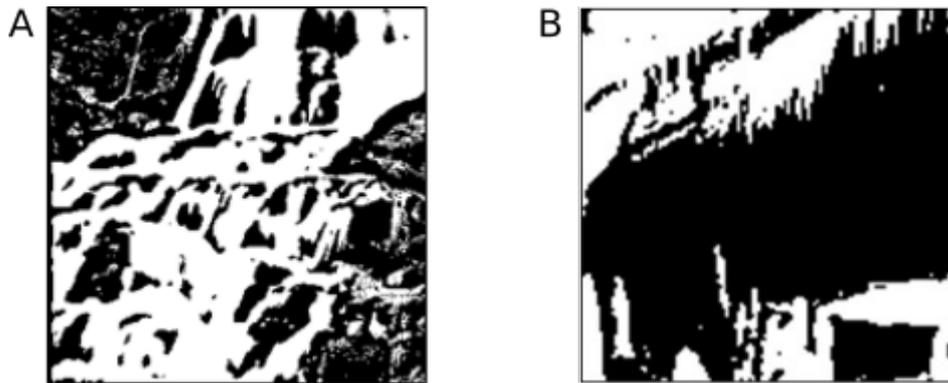


Figure 3: Images manipulated by the author. Originals included in Appendix.



Figure 4: Forrest Scenes.

3.2 Skilled recognition cases

Siegel offers the compelling case of someone who is trained to spot pine trees (2006, p. 492). The subject experiences pine trees in a certain way when they are a novice. Then, as they become an expert pine-spotter, their experience changes. Siegel explains this in terms of the subject coming to perceptually experience pine trees *as pine trees*. An obvious counterpart to the pine-spotter for scene categories would be a skilled spotter of pine *forests*. Imagine that you have been hired to fly a light aircraft over a large area of woodland and to release a fungicide on all and only those forests that are *pine* forests. As you gradually learn how to spot pine forests your perceptual experience of pine forests plausibly changes. Consider the three images in figure 4.

If you were to become an expert pine-forest-spotter, you would recognise immediately that only the first image is a pine forest. Furthermore, this ability would plausibly be associated with a change in your *experience* of the first image. Yet your perceptual experience of the low-level properties of the forest – its shapes and colours – seems to stay the same. I suggest explaining the change in terms of you coming to perceptually experience the scene *as a pine forest*.

A critic might object that pine forest recognition is parasitic on pine tree recognition, meaning that my example is no different to Siegel's object-based contrast case. Given what we know about how scene categorisation works this objection is ill-founded. But to avoid any doubt we should consider some scene categories that do not have such a connection to object categories. Imagine you are a geog-



Figure 5: Desert Scenes.

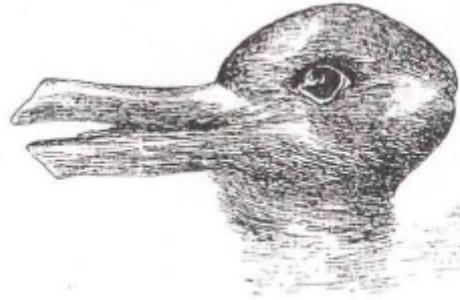


Figure 6: Duck-Rabbit.

rather tasked with mapping the world's *monsoon* deserts. To start with you have to be told which deserts are monsoon deserts as opposed to trade wind deserts etc. Eventually though, you learn to spot monsoon deserts. Gaining this recognitional ability is plausibly reflected in your phenomenology. Consider the three images in figure 5.

Most of us would find it hard to discern which of these is a monsoon desert (it's figure 5C) but if you became an expert desert-spotter it is credible that your experience of the third image would change. Moreover, since there are no objects in this scene, your recognition of this scene category cannot be parasitic on any high-level object categorisation. I suggest explaining this phenomenal contrast in terms of your coming to perceptually experience the scene *as a monsoon desert*.

3.3 Aspect-switching cases

Some of the most vivid and persuasive contrast cases cited by liberals are *aspect-switching* cases. In aspect-switching cases, our phenomenology changes when we switch from seeing an ambiguous image one way to seeing it another way. Figure 6 shows an example that should be familiar.

As you switch between the two ways of seeing this image, your perceptual experience plausibly represents the very same low-level properties. Liberals suggest that the change should be explained in terms of high-level content: you perceptually experience the figure *as a duck* at t_1 then switch to perceptually experiencing it *as a rabbit* at t_2 . I will offer several potential aspect-switching cases involving

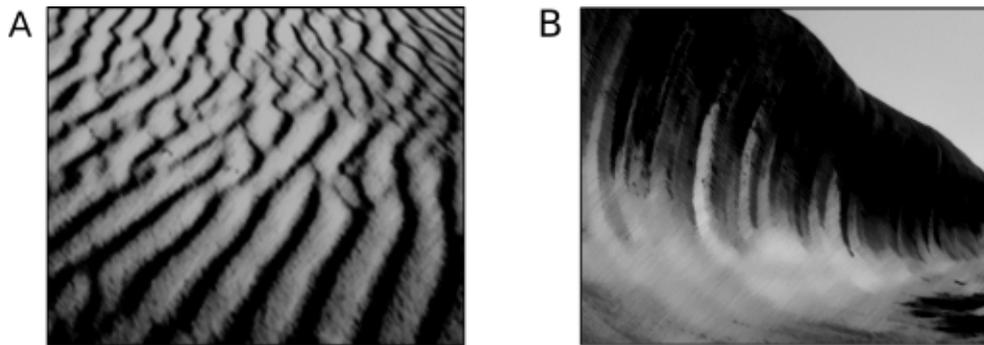


Figure 7: Images manipulated by the author. Originals included in Appendix.



Figure 8: Image manipulated by the author. Original included in Appendix.

visual scenes (informal testing suggests that different cases work for different people). First consider the images in figure 7.

Although both images might be seen as sea scenes neither of them actually is. The first is an image of a desert and the second of a distinctive rock structure in Australia. For each image, you might be able to switch between seeing it *as a sea scene* and seeing it as what it really is. Since the low-level contents of one's perceptual experience seems to stay the same, I suggest explaining this in terms of changes in high-level content: we switch between seeing the depicted scene as being of one category then another. Cases might also be found in which an image is ambiguous between having a certain scene property and having a certain object property. Consider figure 8.

This image might strike you as being of a turtle. In fact, it is a degraded photograph of an island. With this knowledge, you may find it possible to switch between seeing this image *as a turtle* and *as an island*. Where the turtle/island is a scene that looks like a kind of object, figure 9 is a set of objects that look like a kind of scene.

This degraded image could easily be mistaken for a mountain range, but the original image reveals it to be a group of people bent over to form an interlocking pattern of bodies. Again, you might be able to switch between seeing the image in two different ways. Neither of these cases seems to involve changes to the low-



Figure 9: Image manipulated by the author. Original included in Appendix.

level contents of one's perceptual experience, so I would again propose explaining the contrast in terms of high-level perceptual contents: you see the image *as a mountain range* then *as bodies*.

The contrast cases I have offered can be explained in terms of our perceptual experience of high-level properties of visual scenes. But is this the best explanation? To be confident of this we must confront the strongest alternative explanations. The first main alternative is that the phenomenal contrasts can be explained at the level of one's non-perceptual phenomenology. The second main alternative is that the phenomenal contrasts can be explained at the level of one's perceptual attention. I will challenge each of these explanations in the next two sections.

4 Against non-perceptual explanations

One way of resisting my perceptual explanation of the contrast cases is to explain them at the level of the subject's non-perceptual states. It is increasingly thought that one's phenomenology is not exhausted by one's perceptual experience (Bayne & Montague, 2011). If we also enjoy a *cognitive* phenomenology then the contrast cases might be explained in terms of changes in one's cognitive experience.

Price proposes that in the duck-rabbit case "there is some shift in cognitive phenomenal character between thinking of the duck/rabbit figure as a duck and thinking of it as a rabbit." (2009, p. 514) Accordingly, we might describe the experience of the sea/desert in figure 7A in terms of one forming the judgement that the image is a sea scene and then forming the judgement that it is a desert scene. On this view, the judgements are themselves phenomenal states, each with their own distinctive phenomenology. Similar explanations could be offered of the other contrast cases: in the revelation case one comes to judge that the degraded image is a waterfall and in the expert case one gains the ability to immediately judge that the scene before them is a pine forest. This account aligns with my proposal in claiming that the contrasts involve experiencing high-level scene properties. But it diverges from mine in claiming that the relevant experiences are cognitive rather than perceptual.

One way of defending the perceptual account is to deny the existence of cognitive phenomenology. If our phenomenology is exclusively perceptual, then phenomenal contrasts cannot be explained in terms of changes in non-perceptual representations. Although this scepticism is a live option (Prinz, 2012), I think there are good reasons for saying cognitive phenomenology is real (McClelland, 2016). Furthermore, such scepticism would be dialectically problematic here. The reasons for positing cognitive phenomenology closely parallel the reasons for positing high-level perceptual phenomenology (McClelland, 2016), so rejecting the former whilst accepting the latter is a hard position to sustain.⁶ If we countenance both perceptual and cognitive phenomenology, the question is which aspect of our experience accounts for the contrast cases.

It is tempting to argue for the perceptual account on introspective grounds. When I reflect on the phenomenology of the contrast cases, the relevant changes seem to be perceptual rather than cognitive. The difficulty with this move is that introspection is not a reliable guide to whether some content of experience is perceptual or non-perceptual (Block, 2014; Briscoe, 2015; Butterfill, 2009). It is one thing to introspect a change in one's experience but quite another to introspect the nature of the mental state that constitutes that change (Macpherson, 2011, p. 14). If the non-perceptual explanation can be resisted, it must be on other grounds.

A better defence is to show that changes in the contents of one's judgement are neither necessary nor sufficient for the relevant changes in one's phenomenology. In the revelation case, one might judge on the basis of testimony that figure 3B is a cavern and this judgement would entail a particular cognitive phenomenology. Yet your experience would change again when you come to *see* it as a cavern, meaning that your judgement is not the source of the relevant contrast. In the skilled recognition cases, the novice might be in a good position to judge what kind of forest they are flying over. But the relevant phenomenal change comes later when they start to *see* the forests differently. In the aspect-switching cases, I can switch between the two ways of seeing figure 8 whilst maintaining a judgement that it is actually an island. This strongly suggests that making phenomenal judgements about scene categories is not sufficient for having the phenomenal states highlighted by the contrast cases.

Making such judgements is also not necessary for having the highlighted phenomenal states. Siegel (2006) imagines her pine-spotter learning that the forest before them is an elaborate hologram. The subject no longer believes that the objects are pine trees so no longer has whatever cognitive phenomenology is entailed by this belief. Yet the holographic objects still *look* the same way viz. they look like pine trees. Thus the pine-spotter's experience of objects *as pine trees* is perceptual

⁶For the same reasons, accepting cognitive phenomenology whilst reject high-level perceptual phenomenology would be hard to sustain. Conservatives are generally sceptical about *both* of these (Carruthers & Veillet, 2011) so will not be comfortable invoking cognitive phenomenology to rebut liberal accounts of perceptual experience. I take it, though, that one need not be a conservative to resist my account of the contrast cases. One can acknowledge that we perceptually experience high-level properties whilst denying that scene categories are among those properties.

rather than cognitive. By the same token, my expert forest-spotter would cease to believe that the scene beneath them is a pine-forest and yet would still have the same perceptual phenomenology.⁷ Again, the aspect of experience highlighted by the contrast cases comes apart from any cognitive phenomenology entailed by making judgements about scenes.

A critic might respond that the foregoing takes too narrow a view of non-perceptual phenomenology. Perhaps the relevant contrasts can be explained not in terms of judgements or beliefs but in terms of some other kind of non-perceptual representation. Brogaard (2018), for instance, posits ‘phenomenal seemings’ that are neither perceptual nor doxastic. This framework could explain the contrast cases in terms of changes to how things phenomenally seem to one. In the revelation case, for instance, it suddenly phenomenally seems to you that the image is of a cavern. Here the content of your perceptual experience remains constant and the beliefs you form about the image are irrelevant. This alternative explanation should only be taken seriously if we already have good reasons to posit phenomenal seemings (or something like them). As things stand, I do not think this burden has been met (Bayne & McClelland, 2019) so sticking with a binary distinction between perceptual and cognitive phenomenology remains warranted. I hope to have shown in this section that the contrast cases are much better explained in terms of perceiving scene categories rather than in terms of making judgements about them.

5 Against an attentional explanation

Another way of explaining contrast cases is by playing the ‘wild card’ of attention (Fodor, 1988). This strategy has been widely used to resist liberal interpretations of contrast cases (Briscoe, 2015; Carruthers & Veillet, 2011; Price, 2009; Prinz, 2012). The duck-rabbit, for example, is explained in terms of one shifting one’s attention to different parts of the image. This fits well with empirical evidence demonstrating that when subjects see the image the duck way they attend to the ‘beak’ on the left of the and when they see it the rabbit way they attend to the ‘eye’ on the right of the figure (Prinz, 2012, p. 156). Indeed, it is difficult to make the phenomenal switch *without* shifting one’s attention, which presents a serious challenge for liberals.⁸

⁷Thanks to an anonymous referee for suggesting this response.

⁸Fodor appealed to attention to explain the duck-rabbit figure in a slightly different context. Churchland (1988) and others suggested that our perceptual experience changes depending on whether we cognise the image as a duck or a rabbit, thus providing a counterexample to cognitive impenetrability. Fodor (1988) replied that the change is caused by a shift in attention rather than by a cognitive state, allowing cognitive impenetrability to stand. Stokes (2018) shows that Fodor takes too simplistic a view of attention, failing to recognise that the relevant attentional change could itself be a result of one’s cognitive state and so still constitute cognitive penetration. The dialectic in the admissible contents debate is a little different. Liberals and conservatives needn’t argue about what *causes* changes in our attention. Rather, their dispute is over whether such changes exhaust the phenomenal changes in the contrast cases.

Attention is a multi-faceted phenomenon so there are multiple ways in which attention might alter one's perceptual phenomenology. Two visual experiences of the same visual field can differ with respect to how attention is distributed over that field. Our attention can be focused on different *locations*, on different *objects* and on different *properties*. Shifting your attention in any of these three respects can alter your perceptual phenomenology at the level of perceptual *content* and at the level of perceptual *structure*.

At the level of content, attention determines what gets into your perceptual experience. When looking at a visual scene, a small insect might figure in your perceptual experience while you attend to it but cease to do so when your attention moves elsewhere. In principle, a phenomenal contrast might be explained in terms of an attentional shift that alters the contents of your perceptual experience. But this strategy has poor prospects of explaining our contrast cases. The key feature of these cases is that the low-level content of the subject's experience plausibly remains constant across the contrast. Of course, one might argue that attention also alters the *high*-level content of a subject's experience but that would be of no use to conservatives.

At the level of structure, attention imposes a shape on our perceptual field. Perceptual experience typically has a *figure-ground* structure in which an attended item stands out against a background of unattended items. When looking at a bookshelf, for instance, the book on which you focus constitutes an attentional centre relative to which the other books are peripheral. The way that an attended item *pops out* is an important aspect of visual experience. Relatedly, your perceptual experience has a *salience* structure. While you are focussing on a specific region, object or property, things in the periphery can pull on your attention. The salient items draw you toward making *them* the centre of attention, thus forming a new figure-ground structure. So the big shiny book on the left of the shelf might call out for attention while the little dull book on the right does not. Although all three books feature in the contents of your perceptual experience, the *way* they feature depends on these structural aspects of attention – on whether they are in the figure or in the ground and on whether they are high or low in salience.

Appealing to the structural aspects of attention has a lot more promise of accounting for the phenomenal contrast cases. The proposal is that the experience pairs are the same with respect to their content but differ with respect to their attentional structure. The most natural way of unpacking this is in terms of changes to the figure-ground structure of perception. In the original revelation case, the form of the baby suddenly stands out against the background. This is a vivid feature of the phenomenology of revelation cases. In the original skilled observer case, as the pine-spotter becomes more skilled the pine trees plausibly start to pop out from their surroundings. And the specific features diagnostic of being a pine tree – such as the distinctive shape of its needles or colour of its bark – start to pop out too. This fits with empirical work on 'perceptual learning' which suggests that experience in a domain can modulate what we are disposed to attend to in

that domain (Connolly, 2017). And in the original aspect-switching case, we alternate between having the ‘beak’ and the ‘eye’ of the duck-rabbit at the centre of our attentional structure. This fits with the aforementioned studies on the locus of attention during aspect-switching. Whether these attentional differences *exhaust* the phenomenal differences between each experience pair is an open question, but the conservative certainly has an opportunity to argue as much. Can the conservative apply the same explanatory template to the scene-based contrast cases?

Perhaps the phenomenal contrasts can be explained in terms of changes in figure-ground structure caused by a shift in *spatial* attention. However, it doesn’t seem that recognising a scene category is associated with attending to any specific region of the image. You can switch between seeing the sea/desert as a sea scene and a desert scene without changing where you attend. Perhaps the relevant change is brought about by a shift in *object* attention. This is similarly implausible. Most of the images we’re considering don’t even contain objects, and where they do the relevant phenomenal contrast seems to have little to do with how we experience those objects. The point of the contrast cases is that they highlight changes in how we experience *scenes* rather than how we experience objects. One might try to account for the contrasts in terms of the overall scene becoming an object of attention. But the figure-ground structure of perception is contrastive, so where a pine tree can stand out against the background of a forest, a pine forest has nothing against which to stand out. A scene *is* our perceptual background so cannot become the centre of a figure-ground structure.

A more promising approach is to explain the phenomenal contrasts in terms of what *properties* stand out to us. In the revelation case, perhaps the features diagnostic of being a waterfall suddenly pop out for us. In the skilled perception case, as your desert-spotting skills improve perhaps the features distinctive to monsoon deserts pop out to you. And in the aspect-switching case, perhaps you switch between foregrounding the features diagnostic of being a sea scene then foregrounding the features diagnostic of being a desert scene.

But what are the distinctive properties we attend to in these cases? The evidence counts against us attending to localised low-level properties of the scene. Scene recognition is predominately driven by global properties of the overall visual scene rather than by local details within it. Global properties characterise the spatial distribution of boundaries in an overall visual scene and include properties like openness (as illustrated earlier in figure 1), expansion and mean depth. A compelling study by Greene & Oliva (2009) revealed that the global properties of scenes are excellent predictors of how subjects will categorise those scenes, and can even predict the mistakes subjects will make in rapid scene categorisation. Relatedly, a study by Loschky et al. (2019) compared the effects of occluding the foveated region of a scene image and occluding everything but the foveated region of a scene image. They found that the latter frustrated successful scene recognition significantly more than the former, suggesting that scene recognition is driven by peripheral rather than focal perception. Furthermore, a range of evidence sug-

gests that fine-grained local details aren't needed for most scene recognition: we are generally still able to identify the scene category of an image even when it is significantly degraded (Oliva & Torralba, 2006), achromatic (Castelhano & Henderson, 2008) or presented at extreme angles of eccentricity (Boucart et al., 2013). This makes it very unlikely that recognising an image as being of a certain kind of scene involves attending to a specific set of local low-level properties diagnostic of that scene category.

If the contrast cases are going to be explained in terms of changes to figure-ground structures, it had better be *global* low-level properties that become foregrounded when we see a scene as being of a certain category. Assuming that conservatives are happy to countenance these global features as low-level properties, they can explain the contrast cases in terms of attention shifting to features like closedness, low mean depth and moderate expansion. But the main problem for this proposal is that scenes with similar diagnostic global properties can still be experienced with a different perceptual phenomenology. For instance, both sea scenes and desert scenes are associated with the property of openness (Greene & Oliva, 2009, p. 147). But to explain the contrast between seeing the ambiguous image as a sea scene and seeing it as a desert scene, this approach needs different global properties to be foregrounded in each case.

A final option for the conservative is to put aside figure-ground structures and appeal to changes in the salience structure of our perceptual experience. Here the phenomenal contrast is not explained by what's in the foreground of our attention, but by how different aspects of the visual scene call out for our attention. So how might the salience structure of a visual scene change when we put it under a certain category? Here one might appeal to studies on how scene categorisation effects visual search. Torralba et al. (2006) found that scene categories influenced where subjects fixated when searching for objects. Once you recognise a scene as a room, you fixate on the walls rather than the table if you're looking for a painting and fixate on the table not the walls if you're looking for a mug. Recognising a scene category tells us where different kinds of object are likely to be in that scene which in turn tells us where to shift our attention when searching for those objects.

The problem here is that patterns of attention during object-searching do not provide what is needed. The way you search for a person in the sea will be very different to the way you search for a person in a desert, and this will be reflected in different salience structures in your visual experience. But in the contrast cases discussed we are not engaging in visual search at all. So the empirical evidence cited does not give us any reason to think that when you switch to seeing the scene as a desert, the salience structure of your experience thereby changes. Moreover, from a first-person perspective it does not seem that scene categorisation is associated with some distinctive change to one's perceptual salience structure.

So far, the wild card of attention has not given the conservative what they need. Of course, this survey of options does not exhaust the myriad ways in which one might appeal to attention to account for the contrast cases. Crucially though,

I think the empirical literature justifies a *general* scepticism about the prospects of such an attentional account being found. Scene recognition is often achieved without focal attention, which makes it unlikely that the relevant contrast cases can be explained in terms of attentional changes. Perception of scene categories is not subject to inattentional blindness effects (Mack & Rock, 1998), nor is it impaired by subjects performing an attention demanding primary task (Rousselet et al., 2005). Moreover, the high speed at which scene categorisation occurs is thought to be too quick for focal attention to play a role (Biederman et al., 1983). And interesting work by Kietzmann & König (2015) shows how our initial fixations on a novel image are guided by scene categories, suggesting that scene categorisation is achieved prior to focal attention. Cohen et al. (2011) have put pressure on the claim that scene perception can be achieved without *any* attention, but even they acknowledge that the attentional requirements are minimal and that accurate scene perception can be achieved even when focal attention is occupied by an unrelated task. These findings fit with our first-person experience of scene perception. Even when our focal attention is occupied by a particular object, we are nonetheless aware of the kind of scene in which that object is embedded. I don't have to pull my attention away from the cat to see that it is in a garden. Instead, the garden is something I am able to see without focal attention. Given that scene categorisation is achieved without focal attention, it would be surprising if scene categorisation was associated with a distinctive alteration to one's attention.

So although the attention-based approach has some plausibility when applied to object-based contrast cases, its prospects of explaining the scene-based contrast cases are much poorer. With examples like the duck-rabbit, conservatives can cite empirical evidence showing that our attention changes in distinctive ways depending on how we categorise an object. This makes it difficult (though not impossible) for liberals to argue that there is also some *further* change in the contrast cases that involves high-level perceptual contents. But the empirical evidence points in exactly the opposite direction with respect to scene perception: scene categorisation is generally something that's achieved independently of where one focuses one's attention, so an attention-based account of the contrast cases becomes very hard to sustain.

6 Conclusion

I conclude that we perceptually experience the high-level properties of visual scenes. There are plausibly mental representations of scene kinds that satisfy both the Perceptual Condition and the Phenomenal Condition. The case for scene kinds being represented perceptually was primarily driven by empirical data regarding the speed of scene categorisation. The case for perception of scene kinds being phenomenal was driven by a series of phenomenal contrast cases, and the best explanation of these contrast cases is that we perceptually experience the kind properties of visual scenes. Although the contrast cases are driven by first-person

reflection, empirical data played a role in ruling out alternative explanations of the contrasts. More generally, the fact that the empirical data strongly suggests we have perceptual representations of scene categories makes a perceptual interpretation of the contrast cases easier to sustain. Rather than having to posit new mental states to explain the contrasts, the perceptual interpretation can simply suggest that some of the perceptual representations we already have reason to posit are in fact phenomenal representations. The economy of this explanation means that conservatives have their work cut out finding a better explanation. Overall, this division of dialectical labour between third-person and first-person evidence allows us to build a much more robust case for liberalism.

I have discussed what I take to be the two most promising conservative explanations of the contrast cases. However, there are several other explanatory approaches that conservatives might adopt. Conservatives have previously confronted contrast cases by appealing to Gestalt perception (Pautz, 2009), imaginative experiences (Prinz, 2012) and aggregates of low-level properties (Brogaard, 2018). I find it hard to see how these approaches could give a satisfactory explanation of the cases. In fact, as with the attentional approach, scene-based contrast cases seem even *more* resistant to these kinds of explanation than object-based contrast cases. That said, it should be acknowledged that conclusions drawn from an inference to the best explanation might always be overturned by the emergence of a better explanation.

I have argued that we specifically perceive the kind properties of visual scenes. We should also consider what other varieties of high-level scene property might figure in perceptual experience. Do we experience the affordances of scenes, like a forest's property of being navigable in certain ways? This question interacts with emerging empirical work on visual processing of scene affordances (Bonner & Epstein, 2017; Greene et al., 2014). How about the aesthetic properties of scenes such as a landscape's properties of being beautiful, balanced, picturesque or sublime? Could we perceive a painting of this scene as having those properties even though the painting itself is an object? Are these aesthetic properties *kind* properties or *appearance* properties, and what might this mean for our understanding of aesthetic experience (Stokes, 2018)? These questions could interact with valuable empirical work on the distribution of attention in art engagement (Vogt & Magnussen, 2007).

The conclusion that high-level properties of scenes are admissible contents of perceptual experience has a number of promising applications. Regarding phenomenology, it might help us to make sense of how visual experience gives an impression of a rich visual scene despite representing peripheral objects so sparsely. Regarding epistemology, it can help us to understand how our judgements about scenes and contexts are warranted by our perceptual experience. And regarding action, it can help us understand how our perception of an environment guides what we do. So besides adding a new set of properties to the list of admissible contents of perceptual experience, the foregoing offers several promising avenues for future work.

Acknowledgments

Thank you to Tim Bayne and Giulia Martina for comments on an early draft, and to everyone at Institut Jean Nicod for their comments on the talk I gave on this material. This paper was completed with the aid of a Leverhulme Early Career Fellowship and ERC Grant 313552 ‘The Architecture of Consciousness’.

7 APPENDIX: Figure references



Figure 1: Reproduced with permission from [Greene & Oliva \(2010\)](#).



Figure 2: Original and manipulated image reproduced with permission from [Teufel et al. \(2015\)](#).



Figure 3: **A.** Image manipulated by the author using Microsoft Word. Original image from the SUN Database <http://sun.cs.princeton.edu/> (Xiao et al., 2010) **B.** Image manipulated by the author using Microsoft Word. Original image from the SUN Database <http://sun.cs.princeton.edu/> (Xiao et al., 2010).



Figure 4: **A.** Creative Commons photograph by Perseusfs / CC BY-SA https://commons.wikimedia.org/wiki/File:Above_the_pine_forest.jpg
B. Creative Commons photograph by lwtt93 from Shanghai, China / CC BY-SA [https://commons.wikimedia.org/wiki/File:Xinjiang_Fir_forest_\(5211135412\).jpg](https://commons.wikimedia.org/wiki/File:Xinjiang_Fir_forest_(5211135412).jpg)
C. Creative Commons photograph by Johann Jaritz / CC BY-SA https://commons.wikimedia.org/wiki/File:Albeck_Seebachern_Laerchenwald_25102013_869.jpg



Figure 5: **A.** Creative Commons photograph by Fraguando - Own work / CC BY-SA <https://commons.wikimedia.org/w/index.php?curid=78134674>
B. Creative Commons photograph by Fraguando / CC BY-SA https://commons.wikimedia.org/wiki/File:Desierto_del_Sahara.jpg **C.** Creative Commons photograph by Last Emperor / CC BY-SA https://commons.wikimedia.org/wiki/File:Sand_dunes_of_thar_desert.jpg

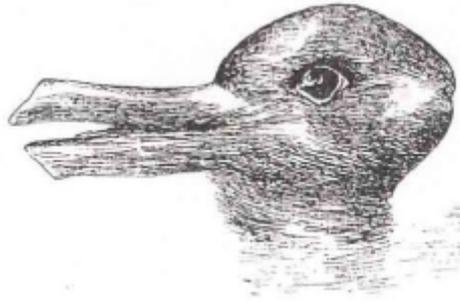


Figure 6: Image in public domain. https://commons.wikimedia.org/wiki/File:Duck-Rabbit_illusion.jpg

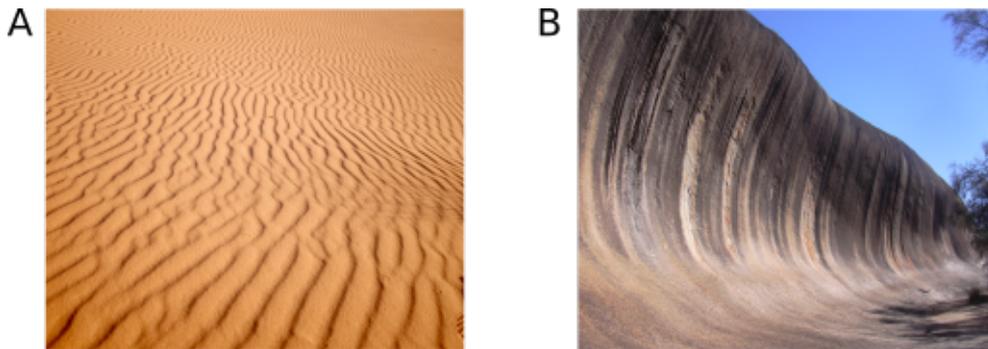


Figure 7: **A.** Image manipulated by the author using Microsoft Word. Original image Creative Commons photograph by Jeremy Cai j / CC0 [https://commons.wikimedia.org/wiki/File:Beach_Waves_In_The_Sand_\(Unsplash\).jpg](https://commons.wikimedia.org/wiki/File:Beach_Waves_In_The_Sand_(Unsplash).jpg) **B.** Image manipulated by the author using Microsoft Word. Original image Creative Commons photograph by Gabriele Delhey / CC BY-SA https://commons.wikimedia.org/wiki/File:Wave_Rock.JPG



Figure 8: Image manipulated by the author using Microsoft Word. Original image taken from <https://www.pickpik.com/island-turtle-krabi-thailand-nature-clouds-77577> under Creative Commons licence.



Figure 9: Image manipulated by the author using Microsoft Word. Original photograph by Carl Warner reproduced with permission by the photographer. <http://www.carlwarner.com/creative/desert-of-sleeping-men/>

References

- Bayne, T. (2009). Perception and the reach of phenomenal content. *The Philosophical Quarterly* (1950-), 59(236), 385–404.
- Bayne, T. (2016). Gist! *Proceedings of the Aristotelian Society*, 116(2), 107–126. <https://doi.org/10.1093/arisoc/aow006>
- Bayne, T., & McClelland, T. (2019). Ensemble representation and the contents of visual experience. *Philosophical Studies*, 176(3), 733–753. <https://doi.org/10.1007/s11098-018-1037-3>
- Bayne, T., & Montague, M. (Eds.). (2011). *Cognitive phenomenology*. Oxford University Press.
- Biederman, I., Teitelbaum, R. C., & Mezzanotte, R. J. (1983). Scene perception: A failure to find a benefit from prior expectancy or familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(3), 411. <https://doi.org/10.1037/0278-7393.9.3.411>
- Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research*, 89(3), 560–572. <https://doi.org/10.1111/phpr.12135>
- Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences*, 114(18), 4793–4798. <https://doi.org/10.1073/pnas.1618228114>
- Boucart, M., Moroni, C., Thibaut, M., Szaffarczyk, S., & Greene, M. (2013). Scene categorization at large visual eccentricities. *Vision Research*, 86, 35–42. <https://doi.org/10.1016/j.visres.2013.04.006>
- Briscoe, R. (2015). Cognitive penetration and the reach of phenomenal content. In A. Raftopoulos & J. Zeimbekis (Eds.), *Cognitive penetrability* (pp. 174–199). Oxford University Press.
- Brogaard, B. (2018). Phenomenal dogmatism, seeming evidentialism and inferential justification. In K. McCain (Ed.), *Believing in accordance with the evidence* (pp. 53–67). Springer.
- Butterfill, S. (2009). Seeing causings and hearing gestures. *The Philosophical Quarterly*, 59(236), 405–428. <https://doi.org/10.1111/j.1467-9213.2008.585.x>
- Carruthers, P., & Veillet, B. (2011). The case against cognitive phenomenology. In T. Bayne & M. Montague (Eds.), *Cognitive phenomenology* (pp. 35–56). Oxford University Press.
- Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology: Human Perception and Performance*, 34(3), 660–675. <https://doi.org/10.1037/0096-1523.34.3.660>
- Churchland, P. M. (1988). Perceptual plasticity and theoretical neutrality: A reply to jerry fodor. *Philosophy of Science*, 55(2), 167–187.
- Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science*, 22(9), 1165–1172. <https://doi.org/10.1177/0956797611419168>
- Connolly, K. (2017). Perceptual learning. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University.
- Epstein, R. A., & Baker, C. I. (2019). Scene perception in the human brain. *Annual Review of Vision Science*, 5(1), 373–397. <https://doi.org/10.1146/annurev-vision-091718-014809>
- Fish, W. (2013). High-level properties and visual experience. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 162(1), 43–55. <https://doi.org/10.1007/s11098-012-9986-4>
- Fodor, J. A. (1988). A reply to churchland's "perceptual plasticity and theoretical neutrality". *Philosophy of Science*, 55(2), 188–198. <https://doi.org/10.1086/289426>
- Greene, M. R., Baldassano, C., Esteva, A., Beck, D. M., & Fei-Fei, L. (2014). Affordances provide a fundamental categorization principle for visual scenes. *arXiv:1411.5340 [Cs, q-Bio]*.
- Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, 20(4), 464–472. <https://doi.org/10.1111/j.1467-9280.2009.02316.x>
- Greene, M. R., & Oliva, A. (2010). High-level aftereffects to global scene properties. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6), 1430–1442. <https://doi.org/10.1037/a0019058>
- Helton, G. (2016). Recent issues in high-level perception. *Philosophy Compass*, 11(12), 851–862. <https://doi.org/10.1111/phc3.12383>
- Kahneman, D. (2011). *Thinking, fast and slow* (Illustrated edition). Macmillan USA.
- Kietzmann, T. C., & König, P. (2015). Effects of contextual information and stimulus ambiguity on overt visual sampling behavior. *Vision Research*, 110, 76–86. <https://doi.org/10.1016/j.visres.2015.02.023>
- Loschky, L. C., Szaffarczyk, S., Beugnet, C., Young, M. E., & Boucart, M. (2019). The contributions of central and peripheral vision to scene-gist recognition with a 180° visual field. *Journal of Vision*, 19(5), 15. <https://doi.org/10.1167/19.5.15>
- Mack, A., & Rock, I. (1998). *Inattention blindness*. The MIT Press.
- Macpherson, F. (2011). Introduction: The admissible contents of experience. In F. Macpherson (Ed.), *The admissible contents of experience*. Wiley-Blackwell.

McClelland, T. (2021). Seeing the forest for the trees: Scene perception and the admissible contents of perceptual experience. *Philosophy and the Mind Sciences*, 2, 2. <https://doi.org/10.33735/phimisci.2021.19>



- Masrour, F. (2011). Is perceptual phenomenology thin? *Philosophy and Phenomenological Research*, 83(2), 366–397. <https://doi.org/10.1111/j.1933-1592.2010.00443.x>
- McClelland, T. (2016). Gappiness and the case for liberalism about phenomenal properties. *The Philosophical Quarterly* (1950-), 66(264), 536–558. <https://doi.org/10.1093/pq/pqv128>
- McClelland, T. (2019). Representing our options: The perception of affordances for bodily and mental action. *Journal of Consciousness Studies*, 26(3-4), 155–180.
- McClelland, T. (2020). The mental affordance hypothesis. *Mind*, 129(514), 401–427. <https://doi.org/10.1093/mind/fzz036>
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3), 353–383. [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3)
- Oliva, A. (2013). Scene perception. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences* (pp. 725–732). MIT.
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, 155, 23–36. [https://doi.org/10.1016/S0079-6123\(06\)55002-2](https://doi.org/10.1016/S0079-6123(06)55002-2)
- Pautz, A. (2009). What are the contents of experiences? *The Philosophical Quarterly*, 59(236), 483–507. <https://doi.org/10.1111/j.1467-9213.2008.584.x>
- Price, R. (2009). Aspect-switching and visual phenomenal character. *The Philosophical Quarterly*, 59(236), 508–518. <https://doi.org/10.1111/j.1467-9213.2009.610.x>
- Prinz, J. (2012). *The conscious brain*. Oxford University Press.
- Rousselet, G., Joubert, O., & Fabre-Thorpe, M. (2005). How long to get to the “gist” of real-world natural scenes? *Visual Cognition*, 12(6), 852–877. <https://doi.org/10.1080/13506280444000553>
- Siegel, S. (2006). Which properties are represented in perception? In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford University Press.
- Siegel, S. (2012). *The contents of visual experience*. Oxford University Press.
- Stokes, D. (2018). Rich perceptual content and aesthetic properties. In A. Bergqvist & R. Cowan (Eds.), *Evaluative perception* (pp. 19–41). Oxford University Press.
- Teufel, C., Subramaniam, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P. R., Goodyer, I. M., & Fletcher, P. C. (2015). Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 112(43), 13401–13406. <https://doi.org/10.1073/pnas.1503916112>
- Torralba, A., Oliva, A., Castelano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, 113(4), 766–786. <https://doi.org/10.1037/0033-295X.113.4.766>
- Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye-movement patterns and visual memory in artists and laymen. *Perception*, 36(1), 91–100. <https://doi.org/10.1068/p5262>
- Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). *SUN database: Large-scale scene recognition from abbey to zoo*. 3485–3492. <https://doi.org/10.1109/CVPR.2010.5539970>

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

McClelland, T. (2021). Seeing the forest for the trees: Scene perception and the admissible contents of perceptual experience. *Philosophy and the Mind Sciences*, 2, 2. <https://doi.org/10.33735/philisci.2021.19>



©The author(s). <https://philosophymindscience.org> ISSN: 2699-0369