Do we have Unconscious Perception?

Craig Stanbury
BIntSt, GC-Arts, GDA-Arts

A thesis submitted for the degree of Master of Arts at
Monash University in 2020
School of Philosophical, Historical and International Studies

---

Copyright notice

© Craig Stanbury (2020).

Abstract

This thesis examines whether humans and animals have unconscious perception. I begin by expounding three cases—blindsight, visual form agnosia, and Continuous Flash Suppression—which all seem to be prima facie instances of unconscious perception. I use these instances to represent the wider case for unconscious perception, and ask: is it beyond a reasonable doubt that each instance is both unconscious *and* perceiving?

I explain that it is an open question whether instances of blindsight and visual form agnosia are unconscious of the stimuli they are allegedly perceiving. We rely on subjective reports to sustain the belief that they are cases of unconscious perception, and subjective reports are an unreliable measure of consciousness (Cheesman & Merikle 1984; Irvine 2013a). I suggest, however, that while there is doubt about whether subjects under the effects of Continuous Flash Suppression are unconscious of stimuli, the doubt is *beyond* a reasonable one. With that said, I then bracket concerns about whether the instances are unconscious and examine whether they are actually perceiving. *If* they are unconscious, are they even instances of perception? I follow the recent trend in unconscious perception debates (e.g. Phillips & Block, 2016) by using Burge's (2010) criteria for perception to frame my analysis. I argue that a critical criterion for Burge—that perception must be attributable to the individual-level—is not satisfied by each instance. Blindsight, visual form agnosia, and Continuous Flash Suppression are *not* individual-level phenomena; therefore, even if we grant that they are unconscious, we have reason to doubt that humans and animals can unconsciously *perceive*.

I finish the thesis by proposing an alternative path to progressing the unconscious perception debate. I suggest that perception is plausibly understood as a manifest kind and argue that *if* it is a manifest kind, then there is room for unconscious perception. Phillips (2018) uses manifest kinds to suggest that there is *no* room for unconscious perception. Therefore, while we agree it is reasonable to hold that perception is a manifest kind, we disagree on what this means for the unconscious perception debate. Indeed, I propose that contrary to Phillips, manifest kinds are an exciting and worthwhile avenue for the unconscious perception advocate to explore. The final position of this thesis, then, is that we *currently* have reasonable doubt whether humans and animals can unconsciously perceive, but there are clear paths forward to removing this doubt. I pave one path, and it is up to others to decide if they want to walk it.

Declaration

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Acknowledgements

A very special thank you to my primary supervisor, Tim Bayne, for the many hours of discussion, guidance, feedback and patience he has provided. Also, thank you to my associate supervisor, Monima Chadha, for her input and encouragement, particularly in the final stages before submission. Thanks to my fellow research student, Niccolò Negro, for supporting me at conferences and for generally being a sounding board to bounce ideas off. Finally, thanks to my partner, Savannah, who believed in me every step of the way.

Table of Contents

List of Abbreviations

2AFC Two alternative forced-choice

AI Artificial Intelligence

CFS Continuous Flash Suppression

b-CFS “breaking” Continuous Flash Suppression

PAS Perceptual Awareness Scale

SDT Signal Detection Theory

Chapter One: The Alleged Instances of Unconscious Perception

Can humans and animals unconsciously perceive? We are not yet in the position to answer 'yes' beyond a reasonable doubt. Explaining why is the topic of Chapters Two and Three. Chapter Four, however, paves a path forward which shows the direction one could take to remove the 'reasonable' doubt. There is elbow room to show that we have unconscious perception.

For now, note that there is widespread acceptance that we *can* unconsciously perceive. As Peters (2017) observes: "most of us in the field do believe the unconscious version of perception exists" (p. 3). Burge (2010), for example, states that perception's "instances are often but not always conscious" (p. 363). Block (in Phillips and Block, 2016), moreover, claims that "seeing is a single fundamental natural kind of which conscious and unconscious seeing are sub-kinds" (p. 169) and that "there is substantial evidence for high level unconscious perception" (p. 178). Blackmore (2004) also holds that "there is no doubt that unconscious learning and perception occur" (p. 280) while Kentridge (2017) states that "if it walks like a duck and quacks like a duck then it is a duck – that unseen primes can be perceived and so that color perception can be unconscious" (p. 5). Kanwisher (2001) further explains that perception is "the extraction and/or representation of perceptual information from a stimulus, without any assumption that such information is necessarily experienced consciously" (p. 90). Examples could go on, but the point is that there seems to be an almost "ubiquitous belief in unconscious perception" (Phillips, 2018, p. 471).

Of course, it is uncontroversial that many mental states "admit of both conscious and non-conscious varieties" (Carruthers, 1996, p.135).[1] Nevertheless, the claim that perception admits of both varieties is a substantial one which deserves special attention. Is perception—as the above quotes indicate—like many other mental states that can be both conscious and unconscious? Or is it a particular type of mental state which can only occur consciously?

To help answer this question, I follow the recent trend in unconscious perception debates by using Burge's criteria for perception as the framework for my analysis.[2] Burge (2010) states

[1] See Berger (2014) for a further general discussion of conscious and unconscious mental states. Also note that Carruthers specifically includes perception as a mental state that comes in both varieties (see pp. 135 - 140).

[2] For examples of how opposing sides of the unconscious debate have staked their claims according to Burge's criteria, see Block (2014, 2016), Phillips & Block (2016), Phillips (2018), and Taylor (2019).

that perception is a natural "psychological kind" (p. 367) which consists of "*objective sensory representation by the individual*" (p. 368) [Burge's emphasis]. For Burge, objectivity involves constancy. So, if we find a mental state which satisfies the following criteria:

- is sensory
- involves constancy
- is individual-level

then according to Burge we have an instance of the natural kind perception. Throughout this thesis, I refer to these criteria as 'Burge's criteria for perception' or simply 'Burge's criteria'. But what do they entail? To be 'sensory' demands that our sense organs gather the information; the perceptual state must begin with sensory stimulation. 'Involves constancy' requires that the sensory state must represent "environmental attributes, or environmental particulars, as the same, despite radically different proximal stimulations" (Burge 2010, p. 114). The 'individual-level' entails that the constancy forming sensory representation must be attributable to the individual and not, as Burge (2010) puts it, merely to a subindividual mechanism "in an individual's visual system or brain" (p. 368).[3]

The type of consciousness the debate is interested in is the "something that it is like" consciousness (see Nagel, 1974). Unless otherwise specified, anytime I mention 'consciousness' in what follows, I am referring to phenomenal consciousness. So, with that said, the question is: are there any mental states which satisfy both Burge's criteria for perception and are unconscious? At first glance, the answer seems to be 'yes'. Consider, for instance, blindsight, DF's visual form agnosia, and Continuous Flash Suppression.

1.1 Blindsight

Blindsight has two hallmarks:

1. Visual capacity in a field defect.
2. No acknowledged awareness of the visual stimulus.

[3] Note that I will often switch between using 'individual/subindividual-level' and 'personal/subpersonal level'. I use 'individual-level' to mirror Burge's (2010) terminology, and 'personal level' to follow Dennett's (1969) terminology when he introduced the distinction. Either way, the two terms are equivalent and are used synonymously.

Blindsight subjects have clinically blind field defects caused by lesions to their primary visual cortex (V1), yet they can still localise and discriminate visual stimuli that they deny seeing (Weiskrantz, 1986; Cowey, 2010).[4]

The first extensively studied blindsight subject, DB (Weiskrantz, Warrington, Sanders & Marshall, 1974), had his right V1 removed due to a tumour. This removal meant that he reported no awareness of visual stimuli in his left visual field (while still being able to accurately report stimuli presented to his normally functioning right visual field). When presented with a striped circle to his blind field, he was asked to guess whether the stripes on the circle were horizontal or vertical. DB had already reported that he could not see the circle—let alone the stripes—yet when pressed to do so, he was able to state the direction of the stripes 90 - 95% of the time. This visual discrimination indicates that even though he could not *consciously* discriminate the stripes, some part of his unconscious brain could. Furthermore, years later, Trevethan, Sahraie and Weiskrantz (2007) discovered that without knowing what the category of the stimulus was, DB could identify outlined figures of animals in his blind field up to 90% of the time (also see Cowey, 2010, for a discussion on this).

Oher blindsight subjects can point towards flashing spots of light despite no reported awareness (Sanders, Warrington, Marshall, & Weiskrantz, 1974). Some subjects can also make saccades and track stimuli that they reportedly cannot see (Stoerig & Cowey, 1997). GY—perhaps the most extensively studied blindsight subject—can detect, localise, and discriminate transient stimuli presented to his blind field (Azzopardi & Cowey, 1998; Azzopardi & Hock, 2011). He can also discriminate between different facial expressions without acknowledged awareness (de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999).[5] In other instances of blindsight, subjects can accurately reach for and grasp targets that they insisted they were not aware of, even if the target was transient (Blythe, Kennard, & Ruddock, 1987; Perenin & Rossetti, 1996). Moreover, Weiskrantz (1997) reports that blindsight subjects can correctly discriminate between colours in their reported blind field. He explains that subjects could "respond to the stimuli that would normally generate the philosophers' favourite species of qualia, namely colours, but in the absence of the very qualia themselves!" (p. 23). As Cowey (2010) states: "It is the absence of visual qualia yet discriminating as if they are present that makes blindsight so interesting" (p. 7).

[4] Here, and throughout the thesis, when I say 'blindsight', I am referring to 'type-1' blindsight which is distinguished from 'type-2' blindsight. 'Type-2' blindsight is when a subject reports a "feeling" that something is present in their blind field but still denies visually perceiving anything in their blind field.

[5] For more examples and review of GY (and thus blindsight in general), see also Kentridge and Heywood (1999), and Kentridge (2015).

Blindsight subjects report that they are not conscious of stimuli in their field defect—so they seem to be unconscious of the stimuli—yet they can often discriminate the very same stimuli at an accuracy rate above 90%, so they seem to be perceiving. Therefore, perhaps blindsight subjects *are* genuine examples of perception without consciousness. For Burge (2010), at least, they fulfil his criteria for perception and are unconscious of the stimuli, he says: "blindsight patients perceive environmental conditions. The perception involves perceptual constancies—including motion, location, and size constancies. The perception guides action. There is strong reason to believe that some of these patients lack phenomenal consciousness in the relevant perceptions" (p. 374).[6]

Blindsight is not the only case to allegedly satisfy both Burge's criteria for perception and be unconscious. Consider Goodale and Milner's (2004) study on DF's visual form agnosia.[7]

1.2 DF's Visual Form Agnosia

Goodale and Milner (2004) explain that due to carbon monoxide poisoning, DF lost many of her conscious visual capacities, yet she retained an ability to discriminate visual stimuli and navigate her surroundings. DF *cannot* see the edges and forms of colours. When looking at a face, for example, she is unable to see the shapes, form and details needed for recognition (p. 9). She also cannot consciously discriminate the difference between a triangle and a square; for her, the shapes run into each other, and so she is unable to report where the triangle ends and the square begins (p. 9). DF also cannot report the angle of a mail slot in front of her; in a series of experiments, she was only at a chance level of success in guessing which angle a slot was (pp. 19 - 21). Nevertheless, when Goodale and Milner asked her to put mail into either a horizontal, vertical or angled slot, they discovered that DF *could* accurately rotate the mail in her hand to enter the slot, no matter the angle.

The same results applied when Goodale and Milner asked her to grasp blocks in variable sizes and orientations (pp. 24 - 26). DF reported having no conscious awareness of the size, shape or orientation of the blocks but could still grasp them correctly. It seems, therefore, that she must have unconsciously discriminated and processed the form and dimensions of

[6] For Burge, 'action guidance' marks the individual-level. I analyse this in detail when I get to Chapter Three.

[7] Note that in their (2004) book Goodale and Milner refer to 'DF' as 'Dee'. However, elsewhere they often refer to her as 'DF', and the literature on her visual form agnosia largely does so too; therefore, I stick with the most commonly used name.

the blocks. These unconscious computations allowed her to scale her grasp's rotation to put her finger and thumb in the correct positions to lift the blocks without any problems.

DF can also hike through a forest while navigating the rough terrain of a rocky path; she can walk the trail and avoid the rocks despite her conscious vision being a host of blurred colours (p. 27). It seems visual information must be available to her in some sort of capacity; DF *has* to be using vision on the hike but, seemingly, it cannot be conscious vision. On the one hand, then, DF exhibits certain behaviour that indicates she can see without problems. She uses visual information about size, orientation and shape to guide her movements. On the other hand, though, DF reports having no conscious awareness of the very visual information that she uses. She cannot consciously report on the information and properties of the objects that guide her computations. How is this possible?

Goodale and Milner (2004) explain that our brains have a dorsal stream, which is used as vision for action, and a ventral stream, which is used as vision for awareness (p. 49). In effect, the dorsal stream is used to guide online, fine-grained motor actions that, in doing so, uses information we are not conscious of. Indeed, scans of DF's brain show evidence that areas of her ventral stream are not working, whereas her dorsal stream is properly functioning (p. 67). This distinction in visual streams reveals why the poisoning that affected DF's visual capacities *only* affected her vision for awareness and not her vision for action. Because of the disrupted ventral stream, DF lost the ability to consciously be aware of the shape and orientation of rocks along her hike. However, because of the intact dorsal stream, she did not lose the ability to calculate the shape and orientation of the rock. Perhaps she was not conscious of the rock, but she was still receiving visual information that enabled her to navigate around it. Goodale and Milner (2004) further clarify:

> What we have learned from these studies is that conscious visual experience of the world is a product of the ventral, not the dorsal stream. You might perceive the tennis ball that has just been lobbed over the net by your opponent but you can never be conscious of the particular information that your visuomotor system uses to guide your successful return. This visuomotor computation happens entirely unconsciously. You are not aware of the fact that the ball is expanding at a certain rate on your retina and that this is an important cue for knowing exactly when to swing to hit it with the sweet spot of the racquet (p. 109).

What the case of DF suggests, then, is that the visual information which we are conscious of forms only a fraction of the visual information our brains use to discriminate external stimuli. As Goodale and Milner (2004) again state:

> We are aware of what one of our visual systems tells us about the world, because we are privy to its products, but there remains a whole realm of visual processing that we can never experience or reflect on, we are certainly aware of the actions that these visuomotor systems control but we have no direct experience of the visual information they use to do so (p. 55).

If our visual system is comprised of a dorsal stream that we may not have conscious access to, perhaps some of the most significant aspects of perception never reach consciousness.

Finally, consider Continuous Flash Suppression (CFS), which Block (in Phillips and Block, 2016) describes as his "favourite case of unconscious seeing" (p. 170).

1.3 Continuous Flash Suppression

CFS exploits the mechanisms that lie behind binocular rivalry, which is a phenomenon that occurs when two images rival each other for our conscious attention, in a winner takes all approach (Koch, 2004). For example, if you were to wear red and green glasses and superimpose an image of, say, a house onto the green lens, and an image of a face onto the red one, then except for brief transitional periods, you would not see an amalgamated picture but instead only see one image at a time (Blake & O'Shea, 2009). The face would remain dominant for a few seconds before the house takes over your conscious vision. If one were not to stop the experiment, then this back and forth would continue *ad infinitum*. It is a striking phenomenon because both eyes are open and staring directly at images superimposed onto one's lenses, yet only one image is present in conscious vision at a time. So, in such a manner, CFS builds upon this mechanism but seems to increase the length and the level of suppression of the unconscious image.

In CFS, an assorted colourful stimulus which rapidly changes up to 10 times per second is presented to one eye (Tsuchiya & Koch, 2005). Called a Mondrian, the stimulus seems to take over the conscious visual experience of both visual fields for anywhere between a few

seconds and three minutes. Whatever stimulus experimenters display to the *non*-Mondrian eye is reported to be consciously suppressed. For example, consider a study by Jiang, Costello, Fang, Huang, and He (2006).

In Jiang et al.'s (2006) study, subjects fixated on a point in front of them with the flashing Mondrian presented to one eye and a nude picture of either a male or female to the other. The Mondrian seemed to suppress conscious awareness of the nude picture reliably; subjects were at a chance level of success in displaying any objective perceptual sensitivity to the stimulus (d' = 0) and insisted that only a Mondrian had been presented to them. Nevertheless, Jiang et al. (2006) discovered that the subject's attention to the nude picture would be attracted or repelled according to their self-reported gender preferences. A heterosexual male's attention, for example, would be attracted to an image of a nude female and repelled by an image of a nude male. Therefore, if the nude picture was reliably suppressed from conscious awareness, then the subjects' attention must have been directed by unconscious processing of the nude stimuli. Similarly, in addition to this Jiang et al. study, consider a "breaking" flash suppression study (b-CFS) by Yang, Zald, and Blake (2007), which also ostensibly shows that our brains can perceive external stimuli in the absence of awareness.

Yang et al.'s (2007) study were on the effects of suppressed facial expressions.[8] The purpose was to understand whether a fearful expression would "break" from suppression to awareness faster than a neutral or happy expression. Consistent with the CFS method, Yang et al. (2007) presented one eye with a flashing Mondrian, and the other with either a fearful, neutral or happy facial expression. Subjects again confidently reported not consciously seeing the non-Mondrian image. Nevertheless, Yang et al. found that the response times for detecting fearful faces were significantly shorter than response times for detecting neutral or happy faces. That is, fearful expressions emerged from suppression to consciousness quicker than a neutral or happy expression.

Both Jiang et al.'s (2006) and Yang et al.'s (2007) CFS studies indicate that stimuli—which show every sign of being suppressed to an unconscious state—can affect our conscious awareness and attention. Even though the nude pictures were not consciously reported, some part of the brain must have discriminated the stimuli. Moreover, even though subjects

[8] Also see Jiang and He's (2006) study which found that the amygdala responds to suppressed faces in a similar way to how it responds to consciously perceived faces. Yang et al's (2007) study was influenced by this study.

seemed to not be consciously discriminating the fearful and neutral facial expressions, some part of their unconscious brain seems to have been. As Blackmore (2004) similarly points out:

> Brain scans have also been used to investigate unconscious perception. For example, an fMRI study showed that unconsciously perceived fearful faces led to greater activity in the amygdala than did happy faces (Whelan et al 1998), and a PET study showed that angry faces activated the right but not the left amygdala (Morris et al 1998). This confirms that even when information is not consciously perceived, it can have effects in the same parts of the brain that would be activated by consciously seen stimuli of the same kind (p. 278).

In other words, CFS studies indicate that our brains can rapidly process, evaluate, and categorise external stimuli before it emerges into consciousness. Furthermore, the quicker emergence into consciousness of fearful faces in Yang et al.'s (2007) study also seems to indicate that the visual processing of the image engaged subjects' individual-level understanding of such expressions. In Jiang et al.'s (2006) study, too, subjects' unconscious response to nudes seemed to reflect their personal level preferences. As Block (in Phillips and Block, 2016) states: "when a sensory registration reflects personal-level *understanding* ... or *values* (Jiang et al., 2006), that is a reason to think it is a personal-level perception" (p. 187) [Block's emphasis].[9] Either way, CFS studies indicate that perception is a mental state which can occur by both partly conscious and unconscious means.[10]

In summary, it seems that the brain can accurately and functionally discriminate stimuli without needing to have any of the conscious awareness that we commonly associate with perception. Blindsight subjects report being consciously blind on one side of their visual field, yet they can use visual information on this blind side to identify, for example, horizontal over vertical lines. In DF's case of visual form agnosia, information that was visually gained, and yet not reported to be conscious, helped her to navigate a hike. In CFS, subjects reported that they could only see a Mondrian, and yet unreported visual information affected their attention. Therefore, blindsight taught us that perhaps our brains have spared visual

[9] In Chapter Three, I argue for a different conception of the personal level.

[10] Just like with blindsight, there are many more CFS studies that could have been canvassed. I have only picked two to stand as a representative sample. For a wider ranging discussion—and defence—of various CFS studies, see Block (2016). Additionally, see Yang et al. (2014) for an in-depth look at the use of CFS in studying visual processing without awareness.

capacities that are unconscious; DF taught us about multiple visual pathways in the brain—one which does not appear to be accessible to consciousness—yet still guides action, and CFS taught us that unreported stimuli can be categorised and affect our attention. As a result, these lessons lend themselves to the widespread belief that we can, indeed, unconsciously perceive.

Let us call blindsight, DF's visual form agnosia, and CFS 'the alleged instances of unconscious perception' or simply the 'alleged instances' for short. I recognise that a full engagement with the case for unconscious perception would require canvassing further instances involving various senses (hearing, etc.). Unfortunately, doing so is beyond my scope. Instead, I will analyse the above alleged instances and trust that they apply *mutatis mutandis* to any other (potential) alleged instance, whether visual or not.

1.4 Moving Forward

On the face of it, the alleged instances of unconscious perception *seem* to be genuine cases of perception without consciousness. However, we need to be careful before yielding the proposition that they *are* genuine cases. Are they unconscious? Are they perception? If so, they should record a 'tick' next to each of the following criteria on our 'unconscious perception report card':

	Blindsight	**DF**	**CFS**
Is unconscious			
Is sensory			
Involves constancy			
Is individual-level			

If no alleged instance records four 'ticks' on our report card, then I suggest there is reasonable doubt whether humans and animals have unconscious perception.

In Chapter Two, I explore whether the alleged instances are unconscious. First, I group blindsight and DF together and argue that it is an open question whether they should record a 'tick' for 'is unconscious'. Both cases rely on subjective measures of consciousness to

sustain the belief that they are unconscious of a stimulus, but, as I explain, such reliance is dubious. Instead of receiving an accurate report on whether a subject is unconscious of a stimulus, we merely receive a report on whether the subject *thought* they were unconscious of a stimulus. Such thoughts (or evaluations) are unreliable. I specifically do *not* claim that blindsight subjects and DF are conscious of the stimuli—that is a much stronger position than what I am committed to—I simply claim that it is an open question either way.

Second, I finish the chapter by briefly discussing whether CFS subjects are unconscious of a stimulus, and grant that in the case of Jiang et al.'s (2006) subjects, it is beyond a reasonable doubt they are. CFS is different from blindsight and DF in that some studies (such as Jiang et al.) do not just rely on subjective measures to claim unconscious perception. As we saw above in section 1.3, Jiang et al.'s subjects are also at a chance level of success in showing objective discrimination of a stimulus. Therefore, I suggest that while there is a possibility Jiang et al.'s subjects are conscious of the nude images under the effects of CFS, it does not seem to be a *live* possibility. Indeed, my analysis will be no mere sceptic's foray into devil's advocacy. There might always be doubt, but I am only interested in *reasonable* doubt.

So, Chapter Two finishes with CFS recording a 'tick' next to 'is unconscious', while it is an open question whether blindsight subjects and DF should:

	Blindsight	**DF**	**CFS**
Is unconscious	**?**	**?**	✓
Is sensory			
Involves constancy			
Is individual-level			

However, despite only one alleged instance recording a 'tick' on our report card, I move forward with analysing whether they satisfy Burge's criteria for perception *as if* each instance is unconscious. In other words, I bracket the concerns from Chapter Two and ask: suppose the alleged instances are unconscious, are they even perceiving?

The alleged instances are sensory. Indeed, there is no debate over whether sensory episodes can be unconscious. Pupil dilation, for example, is sensory, and, as Dretske (2006) points out, allergic reactions carry sensory information. Therefore, we can tick 'is sensory' off our unconscious perception report card:

	Blindsight	DF	CFS
Is unconscious	?	?	✓
Is sensory	✓	✓	✓
Involves constancy			
Is individual-level			

Yet, if perception, as Dretske (2006) again points out, only needs to be sensory, then it would be quite a banal phenomenon and not one that would be of very much interest. In that case, do the alleged instances of unconscious perception involve constancy? While Burge at least thinks blindsight does (see the end of section 1.1. above), there is also evidence to suggest that the alleged instances do not involve constancy (see Phillips 2018, pp. 482-3 for a review). However, even though there is room to move here for either side of the debate, I will not be weighing into the discussion.[11] I do not engage with it for two reasons. One, I do not have the scope to give it an adequate appraisal. Two, my central claim in Chapters Two and Three—that it is *not* beyond a reasonable doubt perception occurs unconsciously—is suitably canvassed without needing to address constancy. Nothing from my main argument is lost by not analysing whether the alleged instances involve constancy. So, because I do not engage with it, we can tick it off the list:

	Blindsight	DF	CFS
Is unconscious	?	?	✓
Is sensory	✓	✓	✓
Involves constancy	✓	✓	✓
Is individual-level			

[11] Although some of what I say about the individual-level in Chapter Three will most likely have implications for 'involves constancy'.

Therefore, we are left with 'is individual-level' as the final hurdle. According to Burge's criteria, for an instance of perception to occur, it must be attributable to the individual-level. In Chapter Three, however, I show that the alleged instances are *not* individual-level phenomena. I begin the chapter by explaining the personal/subpersonal distinction and why it has meaningful ramifications for perception (3.1). I then present my argument for why the alleged instances are not attributable to the personal level (3.2). The argument is as follows:

1. Perception is a personal level phenomenon.
2. Attribution of a personal level state requires the *availability* of the content of a representation for intentional coordination.
3. The content of the representations from the alleged instances of unconscious perception is not available for intentional coordination.

Therefore:

4. The alleged instances of unconscious perception are not instances of perception.

The conclusion is substantial because where one stands on the unconscious perception debate could—if one were to close the question of whether each alleged instance is unconscious—boil down to whether the instances are attributable to the personal level. If I am successful in showing that the above argument is sound, then the individual-level criterion is vitiated and our final analysis of whether perception (according to Burge's criteria) can occur unconsciously, fails to record a 'tick':

	Blindsight	**DF**	**CFS**
Is unconscious	**?**	**?**	✓
Is sensory	✓	✓	✓
Involves constancy	✓	✓	✓
Is individual-level	✗	✗	✗

Therefore, even if we grant that each alleged instance is unconscious, it is not clear that they are perception. More needs to be said (or shown) to move the unconscious perception claim *beyond* a reasonable doubt.

I finish the thesis in Chapter Four, then, by breaking ranks from Burge's criteria and arguing for a different approach to individuating perception, one that paves a path forward in the unconscious perception debate. I argue that intentional coordination—which we encountered in Chapter Three's main argument above—is *so* crucial to perception that instead of being a natural kind, perception is reasonably understood as a manifest kind individuated by the appearance of intentional coordination to the *third-person.*

A manifest kind, as I spell out in 4.1, is a kind that a) has an appearance, and b) is individuated by this appearance. A statue, for example, is a manifest kind because appearing as a statue individuates it from, say, a mere block of stone. Likewise, water vapour is individuated from ice because appearing as fog is *so* crucial to being water vapour that if the constitutive H_20 appeared as ice, we would no longer be referring to water vapour. So, what I put forward in Chapter Four, is that in a similar way to a statue and water vapour, the appearance of intentional coordination individuates perception from non-perception. If this is on the right track—that is, if perception appears to the third-person rather than the first—then it follows that consciousness is not required for perception.

However, Phillips (2018) also suggests perception is a manifest kind, but he proposes that it *does* appear to the first-person via consciousness; thus, there is no room for unconscious perception. So, while Phillips and I both agree that perception is reasonably understood as a manifest kind, we *disagree* on what makes it a manifest kind, and therefore on what the implications for the unconscious perception debate are. My argument in Chapter Four, then, demonstrates that contrary to Phillips, thinking of perception as a manifest kind can show us a way forward to finding unconscious perception.

In sum, the central question in what follows is whether humans and animals have unconscious perception. The answer, as we will see, is 'perhaps': we are not *yet* in a position to answer 'yes' beyond a reasonable doubt. The claim, though, has room to move.

Chapter Two: Are the Alleged Instances Of Unconscious Perception *actually* Unconscious?

This chapter examines the evidence for believing the alleged instances of unconscious perception are unconscious. First, I analyse blindsight and DF and argue that there is reasonable doubt whether they qualify as *unconscious* perception. I begin by spelling out how subjective measures of consciousness sustain the belief they are unconscious of a stimulus. Then, I illustrate how subjective measures do not reliably distinguish between unconscious processing and *degraded* conscious information. Therefore, I conclude that it is a live possibility blindsight subjects and DF are conscious of a stimulus but fail to report this consciousness.

Next, I briefly analyse CFS. I suggest that insofar as CFS does not just rely on subjective measures to claim stimulus suppression, the problems it faces in claiming *unconscious* perception are less severe than the ones faced by blindsight and DF. There is doubt about CFS, but in some cases—particularly for Jiang et al.'s (2006) study—the doubt seems to be beyond a reasonable one.

2.1 The Problems for Blindsight Subjects and DF's Visual Form Agnosia

To be justified in believing that an alleged instance of unconscious perception is actually unconscious, our belief must be sustained by a reliable measure. A reliable measure of consciousness should be as close to being both exhaustive and exclusive as possible.[12] Exhaustiveness requires that the measure is sensitive to *all* relevant aspects of conscious experience; it should not characterise conscious behaviour as unconscious mental processes. Exclusivity, on the other hand, demands that the measure should *only* measure conscious experience; it should avoid characterising unconscious processes as conscious ones (Reingold & Merikle, 1990; Overgaard, 2015).[13] It is unlikely that any measure of consciousness will ever be *completely* exhaustive and exclusive; infallibility is an unrealistic

[12] This standard for epistemic success is similar to—and influenced by—reliabilist epistemology, which holds that: Person S is justified in believing a proposition P if and only if S's belief that P is produced or sustained by a process M and M is reliable (see Goldman, 1979).

[13] Another way of putting the demands for reliability are via Irvine's (2019) criteria for a reliable, valid and accurate measurement. She explains that the measurement must be stable enough for repeatability and replication. It must measure the target it is meant to and not a property that correlates with the target. And it must accurately reflect the phenomenon being measured with minimum error. I take these demands to be mostly captured by being both exhaustive and exclusive.

standard. Instead, the closer a measure gets to infallibility, the more reliable it becomes. The cases of blindsight and DF, however, rely on subjective measures of consciousness to sustain the belief they are unconscious of a stimulus, and subjective measures are unreliable.

A subjective measure of consciousness regards someone as conscious of a stimulus if they report consciousness of that stimulus. The report forms a threshold which is supposed to mark the distinction between conscious and unconscious perception. If a positive report is given, then the stimulus falls above the threshold, and if a negative report is given, then it falls below the threshold. The standard tool used for generating reports is introspection: the first-person judgement of current and ongoing mental states (Overgaard & Sandberg, 2012, and see Schwitzgebel, 2019). It is helpful to compare subjective measures to objective measures.

An objective measure of consciousness considers someone as conscious of a stimulus if they show an appropriate behavioural response towards the stimulus. According to objective measures, if a subject displays greater than zero discrimination of a stimulus (d' > 0), then they fall above the objective threshold and are thus most likely conscious of the stimulus. If they have no behavioural response towards a stimulus (d' = 0), or their responses are at chance, then they fall below the objective threshold, indicating that they are *not* conscious of the stimulus (see Seth, Dienes, Cleeremans, Overgaard, and Pessoa, 2008, and Cheesman and Merikle, 1984, 1986). In most scenarios, the objective and subjective thresholds seem to align. When you report seeing a red strawberry, you can also show a behavioural response to the strawberry. Blindsight and DF, however, are peculiar cases because the thresholds dissociate: stimuli fall above the objective threshold but below the subjective. An advocate of these cases being unconscious perception, therefore, cannot endorse objective measures as reliably exclusive otherwise the cases are regarded as conscious by default. Instead, they must pin their hopes to subjective measures being reliable.

The claim from the advocate is that (especially in these peculiar cases) the objective threshold simply tracks perceptual sensitivity without tracking consciousness. In other words, while d' > 0 *might* be an exhaustive measure of consciousness (Irvine, 2013a, 2013b), blindsight, for instance, demonstrates that it is not an exclusive measure. As Weiskrantz (1997) explains: "a behavioural response to a stimulus may be necessary but it is certainly not sufficient for attributing conscious awareness of that stimulus to the responder" (p. 80 -

see also Frith, Perry, and Lumer, 1999).[14] So, the upshot, as the claim goes, is that if a stimulus falls between thresholds, we have good reason for thinking there is an instance of unconscious perception; that is, we have an instance of perceptual sensitivity in the *absence* of acknowledged awareness.[15] Therefore, the belief that a blindsight subject or DF is unconscious of a stimulus hinges on them giving a negative subjective report. As Cowey (2010) states: "The 'blind' in blindsight depends on the patients' verbally reporting 'no' when they are asked if they saw a visual stimulus" (p. 4). However, a negative subjective report can only sustain the belief that we have unconscious perception if subjective reports are reliably exhaustive.

The purpose in what follows is to show that subjective measures are *not* reliably exhaustive, particularly when trying to capture degraded conscious information.[16] This lack of reliability is a serious concern and germane to our topic. If one wishes to claim that blindsight subjects and DF are instances of unconscious perception, then one must precisely rule out—beyond a reasonable doubt—that the reason for the dissociation between thresholds is that subjects are actually experiencing low and degraded conscious information. But because of the unreliability in subjective reports, it is a live possibility that blindsight and DF experience the "visual equivalent of the faintest of murmurs … [seeing] through a glass, darkly" (Phillips, 2016, p. 441).[17]

2.1.1 Subjective Measures and the Problem of the Criterion

As mentioned in Chapter One, and as many before me have pointed out, when gathering a subjective report, instead of obtaining an accurate report on whether subjects are conscious of a stimulus, we merely obtain a report on whether a subject *thought* they were conscious of

[14] For more on the basic problem of assuming that behaviour marks consciousness, and for discussion in trying to operationalise *phenomenal* consciousness as behavioural abilities rather than a subjective "what is it like" experience, see, for example, Spener (forthcoming), Timmermans and Cleeremans (2015), Overgaard (2006), and in particular Irvine (2013a).

[15] I purposely stop short of endorsing d' > 0 as perception rather than mere sensory discrimination. As we will see in Chapter Three, d' > 0 does not necessarily nor sufficiently meet the demands for perception. Recall that this chapter brackets the concern over whether the alleged instances are perception, and asks *if* they are perception, are the subjects even unconscious?

[16] For discussion on situations when introspection could be reliable see, for example, Bayne and Spener (2010), Kriegel (2013), Spener (2015), and Wu (2018).

[17] I do not dedicate too much space in this chapter to *directly* applying the issues with subjective measures onto DF. This is because I use any discussion on blindsight to equally apply to DF. Indeed, the issues faced by both blindsight and DF throughout this chapter are so similar that if I were to keep directly applying the problems of subjective measures on to both, I would repeat myself almost ad verbatim.

a stimulus, and such thoughts (or evaluations) are prone to error. Schwitzgebel (2008), for example, remarks that "the introspection of current conscious experience, far from being secure, nearly infallible, is faulty, untrustworthy, and misleading—not just *possibly* mistaken, but massively and pervasively" (p. 259) [Schwitzgebel's emphasis]. Snodgrass and Lepisto (2007) further state that "there is no such thing as an unmediated 'subjective report' - ever" (p. 526), and Kentridge (2015) iterates that "it is dangerous to rely on introspection" (p. 41). Irvine (2009), moreover, asserts that "subjective reports can be very unreliable measures of what a subject is experiencing" (p. 553), and Cheesman and Merikle (1984) conclude that "by itself, the subjective threshold is a useless measure of awareness". Hohwy (2011) labels this scepticism about the reliability of introspection the 'Argument from Variability':

> "AV1 There is evidence of introspective variability across conditions and across subjects.
> AV2 Introspective variability across conditions and subjects is best explained by introspection's being unreliable.
> So, by inference to the best explanation, introspection is unreliable" (p. 265).

In what follows, I explore this unreliability by first using the problems with Signal Detection Theory (SDT) to represent the problems with subjective measures of consciousness more generally. Then, I analyse the exclusion failure paradigm and show that rather than solving the problems which SDT poses, it extends them. Thirdly, I examine the Perceptual Awareness Scale to a) show that whether a subject acknowledges awareness or not can hinge on their task instructions rather than the presence or absence of consciousness, and b) emphasise the dangers of making a measure *so* exhaustive that it precludes exclusivity. Because the cases of blindsight and DF rely on subjective measures to sustain the belief they are unconscious of a stimulus, each issue canvassed below *directly* applies to them.

SDT seeks to determine at which point a signal (or stimulus) is strong enough for detection (Green & Swets 1966). The way it does so is by measuring both a subject's objective perceptual sensitivity to a stimulus and by gathering a subjective report. Perceptual sensitivity is usually measured via a two-alternative forced-choice task (2afc). 2afc tasks a subject with saying which one of two intervals a stimulus was presented in. For example, in an Azzopardi and Cowey (1998) study, subject GY was presented with two temporal intervals with a target stimulus placed in only one. Insofar as GY picked the correct interval across a sufficient number of 2afc tasks, he displayed perceptual sensitivity to the stimulus

(d' > 0).[18] By forcing subjects to choose between two intervals, 2afc ensures that subjects adopt a symmetrical criterion which simply chooses the strongest sensory stimulation, and is thus naturally unbiased. Cheesman and Merikle (1984) explain: "This method ... has the advantage of removing the no or negative response from the set of possible answers. Individual differences in the willingness to guess a response were thus avoided" (p. 388).

Complications arise, however, when subjects are asked to provide a yes/no report on whether they were aware of the stimulus. A yes/no response is different from 2afc, because under the yes/no paradigm, subjects are "free" to shift their response criterion according to independent variables. Cowey (2010) explains: "A subject's judgement in a yes/no task depends not only on sensitivity to the target, but also on any tendency to select one or other response independently of his sensitivity, i.e. response bias" (p. 9). In the same Azzopardi and Cowey (1998) study, for instance, GY was also asked to respond 'yes' if he saw something and 'no' if he did not; they found, however, that his yes/no report *dissociated* from his perceptual sensitivity. Is this dissociation between thresholds due to perceptual sensitivity in the absence of consciousness? Or is it due to degraded conscious information which is not positively reported? While I think these questions remain somewhat open, there is significant evidence to suggest that the dissociation in blindsight is due to a bias criterion. To spell out the problem of the criterion further, consider the four possible response pairings in a yes/no task:

	Yes	No
Present	Hit	Miss
Absent	False Alarm	Correct Rejection

If a stimulus signal, say, a green light, is present and a subject says 'yes' that they can see the green light, then they record a 'hit'. However, if the subject says 'no' when the green light is present, then they record a 'miss'. If *no* green light was presented, though, and they respond 'yes', then they would record a 'false alarm', and if they respond 'no' with no green light present, then they record a 'correct rejection'.

[18] Other 2afc trials can be spatial instead of temporal; for example, forcing a subject to say whether a stimulus was presented in the top or bottom half of a screen.

What's more, two subjects who are under the same stimulus conditions, and exhibiting the same objective sensitivity, can differ in their subjective reports. One subject might be cautious, so they apply a conservative criterion which results in a high threshold. Perhaps they are motivated by minimising 'false alarms'. The other subject, however, might be rash so applies a liberal criterion, resulting in a low threshold. Perhaps they are motivated by maximising 'hits' and so are not worried about 'false alarms' (see Spener, forthcoming).

The benefit of a conservative criterion is that the chances of recording 'false alarms' are small. The disadvantage is that it increases the chances of recording a 'miss'. Therefore, a conservative criterion can often fail to be exhaustive: it does not capture *all* instances of consciousness. If the green light is dim or fleeting, and therefore the conscious information of the light is degraded, it is highly likely that a subject with a conservative criterion will 'miss' reporting conscious awareness. Some conscious information might be so degraded that it goes unreported—not because it was not there—but because it did not meet the subject's criterion for responding 'yes' (see Eriksen, 1960).

However, adopting a liberal criterion does not solve the problem. While it has the advantage that the chances of recording a 'miss' are low, it comes with the added cost that the chances of recording a 'false alarm' increase. Subjects with a liberal criterion are susceptible to overstating the extent of conscious awareness; thus, their subjective reports fail to be exclusive: they do not capture *only* instances of consciousness. Therefore, we cannot solve the exhaustiveness problem by making the criterion as liberal as possible (Overgaard and Sandberg, 2012, and see 2.1.3 below).

The point is that whether a subject responds 'yes' or 'no' hinges on their response criterion, and not on whether they were aware of the stimulus *as such*. Irvine (2009) explains: "since this threshold is liable to change with the criterion level, it does not accurately reflect the subject's underlying ability to discriminate or identify stimuli" (p. 553). Eriksen (1960) also elaborates: "these marked individual differences indicate that as long as we rely upon the *S* [subject] himself to report whether or not a verbal judgment was made with awareness, we are in fact using as many different criteria of awareness as we have experimental *S*s [subjects]" (pp. 292-3). Furthermore, whether a subject responds 'yes' or 'no' (that is, has a conservative or liberal criterion) can hinge on trivial variables independent of conscious awareness.

If an experimenter emphasises, for example, the importance of not recording any 'false alarms', it is likely that the subject will adopt a conservative criterion. When a subject is only paid, say, $1 for every time they record a 'hit' but are fined $100 for every 'false alarm', their criterion will be conservative. Likewise, if a subject receives $100 for a 'hit' but is only penalised $1 for a 'false alarm', he will adopt a liberal criterion (see Merikle, 1984). While this example is an extreme one, the bottom line is that it is hard to imagine that such fluctuations in a criterion due to monetary incentives reliably track fluctuations in *actual* conscious awareness. Moreover, *other* variables such as task type, difficulty level, motivation, expectations and confidence can all affect a subject's criterion for reporting 'yes' or 'no'. For instance, consider perceptual defence, which gives us an example of how expectations and confidence can compromise a report.

Perceptual defence occurs when a subject is shown two words under identical stimulus conditions but are less likely to record a 'hit' for one word simply due to a bias against that word. The first word, for example, might be inoffensive, such as 'shot', whereas the second word might be sexually explicit or a swear word, such as 'shit'. When subjects report awareness of 'shot' more often than they do for 'shit', they exhibit perceptual defence towards that word (Irvine, 2009, 2013a). Subjects may not have expected to detect a swear word; they may have felt embarrassed, taken aback, or dismissive that such a stimulus could ever be presented. Therefore, before responding 'yes', they needed to have a high degree of confidence that they saw the word (whereas they only needed a low or moderate degree of confidence to respond 'yes' for a neutral word). With a stimulus variable of only one letter, it does seem that *the* variable which determines whether the swear word was reported or not is how confident the subject was.

If this is the case, though, then the upshot is that conscious experience—the very phenomenon that a subject's report is supposed to distinguish—is not sufficient for a positive report; sometimes subjects need to be both conscious *and* confident before reporting 'yes'. Cheesman and Merikle (1984) explain:

> It is not possible to determine if the absence of discriminative responding is due to a lack of perceptual sensitivity to the presence of a stimulus or to a lack of confidence that a particular stimulus has been presented ... On the other hand, the data from the subjective threshold condition indicate that a

> considerable amount of perceptual processing occurs when subjects lack confidence that a stimulus has been presented (pp. 394-5).

These considerations raise reasonable doubt as to whether subjective reports are exhaustive. Indeed, they raise reasonable doubt as to whether blindsight subjects and DF are *actually* unconscious of a stimulus when they say they are.

Perhaps conscious information is available to blindsight subjects and DF, but the information is *so* dim and degraded that they lack the required confidence for a positive report. The blindsight subject, for example, might have expected their conscious information to be rich and pronounced—the equivalent, say, of a *bright* green light—but because the information was so dim and degraded—similar to, say, a *faint* green light—they were not confident enough to report 'yes'; the degraded information did not meet their criterion for a 'hit'. As Dennett (1991) points out: "if he [the blindsight subject] complains of the absence of qualia, he might simply be noticing the relative paucity of information he now gets from his vision and misdescribing it" (pp. 358–9). Consider again the two hallmarks of blindsight:

1. Visual capacity in a field defect.
2. No acknowledged awareness of the visual stimulus.

But because of the problem of the criterion, a subject can satisfy (2) and yet still be conscious of the stimulus. Indeed, it seems that whether a subject satisfies (2) can hinge on variables independent of the presence or absence of consciousness.[19] Therefore, it is a live possibility that blindsight is the result of degraded conscious information combined with a variable response criterion (see Campion, Latto, & Smith, 1983). Phillips (2018) makes a similar point:

> A failure to report a stimulus in a biased yn [yes/no] task together with the presence of perceptual sensitivity (d' > 0) does not suffice to show unconscious perception. It is perfectly consistent with the subject being conscious (albeit perhaps in a dim, distorted or degraded manner) but operating with a conservative response criterion in relation to her conscious experience. Such a possibility is not outlandish (p. 485).

[19] As Azzopardi and Cowey (1998) point out: "the degree to which the subject is susceptible to encouragement and the extent of any encouragement that is given could therefore determine whether or not he would be judged to have blindsight" (p. 304).

Can we *reasonably* rule this out? I do not think so. Moreover, it is plausible that the visual information DF uses to navigate her surroundings is conscious, yet is so dim and degraded that it falls below her criterion for responding 'yes'. Perhaps DF's conscious visual information that she *does* acknowledge experiencing (blurred colours, etc.), is quite rich and pronounced, so any conscious information that is severely degraded fails to be positively reported. Either way, whenever there is a dissociation between thresholds, the dissociation can only sustain the belief that we have unconscious perception if we know—beyond a reasonable doubt—that a subjective report is a reliably exhaustive measure of consciousness. As we have just seen, however, consciousness is not reliably present *only* when a subject responds 'yes'. For this reason, perhaps we should reject any procedure which relies solely on subjective reports to distinguish between unconscious information processing and degraded conscious perception (see, for instance, Reingold and Merikle 1990, pp. 17-18, and Holender, 1986). Perhaps a different measure that does not depend on such a variable criterion is needed.

One method that has been put forward to solve the problem of capturing degraded conscious information is the 'exclusion failure paradigm' (Debner & Jacoby, 1994; Merikle & Joordens, 1997). As we are about to see, though, despite initial appearances, it is also vulnerable to the same problems which SDT faces. Indeed, my aim in addressing the exclusion failure paradigm is to deepen our understanding of how a criterion can be variable according to factors independent of consciousness. We will further see that if a subjective report is not *just* sensitive to conscious information, it is an open question whether it has captured all and only instances of consciousness.

2.1.2 The Exclusion Failure Paradigm

The exclusion failure paradigm regards a subject's ability to exclude a briefly presented target word as the qualitative difference between conscious and unconscious perception. A subject is given a target word, say, 'leader', and are then shown the word stem 'lea' for a duration of 50ms. The subject is instructed to complete the word stem with a word that is *not* 'leader'; if they follow the task instructions successfully, they are regarded as conscious of the word stem. However, if they are unable to follow the task instructions, and thus fail to exclude 'leader', they are regarded as *not* conscious of the word stem.

The extra cognition, control, and concentration it takes to overcome one's primes and successfully exclude the target word is only possible, as the reasoning goes, when one is conscious of the stimulus. The *conscious* ability to control our responses leads to exclusion success. The inability, therefore, to follow task instructions and overcome our primes indicates automatic and unconscious perception of a stimulus. Exclusion success or failure does seem to be an auspicious method for distinguishing between conscious and unconscious perception. Could it be *the* reliable measure we are looking for? Not quite.

Instead of marking a distinction between conscious and unconscious perception, exclusion success is variable according to one's bias criterion. For instance, consider a Visser and Merikle (1999) study which investigated the effects of monetary incentives on exclusion success. In the study, half of their subjects were shown a primed word stem at 250ms, and they were able to exclude the word successfully. The same subjects, however, were then shown the word stem at 50ms, and were *not* able to exclude the target word. According to the exclusion failure paradigm logic, subjects were conscious of the word at 250ms but not at 50ms.

The *other* half of their subjects, though, were presented with the word stem under the same stimulus conditions but with the added variable of a monetary incentive. These subjects started with $15 and were penalised $1 for every error made; naturally, this encouraged them to follow task instructions as carefully as possible to minimise exclusion failure. Indeed, with the extra motivation, Visser and Merikle observed substantial improvement in the subjects' exclusion success. What should we make of this? Did subjects all of a sudden become conscious of the word stem with increased motivation? No. Instead, because subjects' motivation levels changed, their response criterion for reporting awareness changed. If exclusion failure genuinely indicated unconscious perception, then adopting a more liberal criterion should not influence results (see Snodgrass, 2002).

Block (2005) argues that the reason subjects could not report the stimulus without a monetary incentive was that phenomenal consciousness had *overflowed* cognitive access. Hence, when they could exclude the target word with increased motivation, it was because the task conditions enabled the overflowed consciousness to become accessible for a positive report. Notwithstanding the substantial debate surrounding overflow consciousness, insofar as Block's interpretation is correct, it causes trouble for subjective measures of consciousness. If there is *any* overflow, then it follows that subjective measures fail to be

exhaustive.[20] Nevertheless, Irvine (2009) offers an alternative explanation. She argues that instead of the conscious information being inaccessible, the information may have just been *so* consciously degraded that it failed to meet subjects' criterion for exclusion under the specific task conditions. She cites a Fisk and Haase (2006) study to help make her point.[21]

Fisk and Haase (2006) altered the exclusion failure paradigm to a forced-choice report. In doing so, they discovered that exclusion failure could simply be due to task difficulty. In their study, subjects were initially shown a previously primed word stem for a duration of 50ms. Without a monetary incentive, the subjects were again asked to provide a free-choice report, but again they failed to exclude the target word. However, when the subjects were given a forced-choice between two words ('place' and 'night'), and instructed to choose the word which was *not* the target word, subjects were able to overcome their primes and achieve exclusion success. If exclusion is only possible when using conscious information, it seems that subjects were conscious of the word all along; they just needed an easier task to make use of the degraded conscious information. For our purposes, there are a few lessons to take from these studies.

First, it is likely that the reason subjects could not report the stimulus in the free-choice paradigm was not that the information was inaccessible or unconscious, but because the conscious information was degraded (Irvine, 2009). By merely altering the task difficulty, subjects were able to use the *same* amount of conscious information from the free-choice response to successfully exclude the target word. Therefore, a subject's response criterion can be variable even when the available conscious information is not. Falling below a subject's criterion for reporting awareness does not conclusively tell us that a person is not conscious of a stimulus, it only tells us that under the particular task type or difficulty level it did not meet their *criterion* for reporting awareness. As Timmermans and Cleeremans (2015)

[20] For more on overflow consciousness, see Block (2007, 2011), Sperling (1960), and Sligte et al. (2008). See also Phillips (2016) for a discussion on the dilemma Block's commitment to overflow consciousness poses for him in claiming the alleged instances of unconscious perception are unconscious. However, also see D'Aloisio-Montilla (2019) for an analysis on the problems it poses for both Block *and* Phillips. Note that for different reasons, I address overflow consciousness again in Chapter Three.

[21] For a further study supporting the conclusion that exclusion success depends on a subject's criterion see Bengson and Hutchison (2007). They also tested for variability in response criteria under exclusion tasks and concluded that "the ability to act upon (via exclusion performance) and report information in a masked prime is determined by a variable response criterion, which can be manipulated as an independent variable" (p. 785). Also see Sandberg et al. (2014) who used the Perceptual Awareness Scale to reach the same conclusion, and see Snodgrass (2002) for a longer argument that exclusion failure is susceptible to the problems in SDT.

explain: “one cannot emphasize enough how apparently small differences in procedures may lead one to strikingly different conclusions when it comes to distinguishing between conscious and unconscious cognition” (p. 41).

Moreover, if a criterion is susceptible to independent variables such as monetary incentives and task difficulty, it is hard to determine where the *actual* threshold between conscious and unconscious processing is. If the subjective threshold is not just indexed to the presence of consciousness, we cannot claim beyond a reasonable doubt that all instances of (degraded) conscious information are accounted for. It is an open question whether some conscious information is just *too* degraded for exclusion success (or a positive report).

However, are there any measures sensitive enough to reliably capture this degraded conscious information? Perhaps. But by considering the Perceptual Awareness Scale (PAS) we will further see the dangers of trying to solve the exhaustiveness problem by making a measure exceedingly sensitive to consciousness.

2.1.3 The Perceptual Awareness Scale and More Problems

Rather than a binary answer, PAS offers subjects four different scale points to choose from (Ramsøy & Overgaard, 2004):

- Clear image
- Almost clear image
- Weak glimpse
- Not seen

A response paradigm more nuanced than a yes/no report is potentially helpful for distinguishing between degraded conscious information and unconscious visual processing. Consider, for instance, Overgaard, Fehl, Mouridsen, Bergholt, and Cleeremans (2008) study on the ostensibly blindsight subject GR. By utilising the PAS scale, Overgaard et al. (2008) elicited the presence of degraded conscious information that was not reported in a yes/no task.[22]

[22] Note, however, that I use this study as a representative of PAS studies in general, and so the subsequent analysis of its problems likewise applies to the use of PAS in *any* of the alleged instances of unconscious perception.

GR initially exhibited both hallmarks of blindsight: she had visual capabilities in a field defect and responded ‘no’ when asked whether she was aware of the stimulus. However, when Overgaard et al. (2008) applied the PAS scale to GR under identical stimulus conditions as the yes/no paradigm, she reported awareness of the stimulus somewhere between ‘almost clear image’ and ‘weak glimpse’. The reason she was judged to have blindsight in the first place, therefore, was not that she lacked a conscious experience, but because her criterion for reporting the experience was too conservative. Overgaard et al.’s (2008) study demonstrated that under a more sensitive measure, GR’s ability to discriminate stimuli does not, as they put it, “reflect unconscious vision, but rather degraded, yet conscious vision” (p. 4). Indeed, aside from anything else, PAS gives us another good example of how instead of hinging on the presence or absence of consciousness, determining whether a subject has blindsight or not can be variable according to task instructions.

However, even though PAS offers compelling evidence to infer that blindsight subjects are conscious of the degraded information, it is still not clear whether they are or not. Firstly, and most importantly, it is relevant to question whether PAS is an exclusive measure of consciousness. Recall that the more sensitive a measure is, the higher the risk that subjects will confuse unconscious information processing with conscious information. As Draines and Seth (2010) explain:

> If one saw a square but was only aware of seeing a flash of something, then one has not consciously seen a square … Perception of the fact it was a square could be entirely unconscious even as participants knew (with varying clarity as measured by PAS) that something was presented. But of course, participants in these experiments will always be aware of seeing something in the world (unless they blink), even if they think they are seeing a blank screen; the question is not whether they were aware of seeing something was present but whether the content we can infer from discrimination performance was conscious (pp. 1079 - 80).

In other words, claiming that an especially sensitive measure (like PAS) indicates conscious perception, is only viable if we can ensure that the measure is also exclusive; if not, we may *preclude* unconscious perception.

Secondly, it also seems that PAS is particularly susceptible to variations in confidence (Irvine, 2013b). When a subject is evaluating which point on the scale best reflects their experience, their confidence levels likely influence their decision. The difference between choosing a 'weak glimpse' or 'not seen' could simply depend on how confident they were that they experienced the degraded information, and not on whether the stimulus was *actually* experienced. Finally, if PAS is intended to measure phenomenal clarity, it is unclear how it can be used as a transferable metric across subjects and experiments (Irvine, 2013a, 2013b). How do we verify if one person's clarity is the same as another's? That is, how do we know whether subject A's 'weak glimpse' is the equivalent to subject B's? These questions lead us into a meta-problem which applies to *any* subjective measure; namely: without knowing what the contents of consciousness actually are, we have no independent measure with which to validate a subjective report.

2.1.4 No Independent Verification

Not having an independent measure is a serious problem because it means that we cannot verify how exhaustive or exclusive a subjective report is. Perhaps if we knew, for instance, whether (or to what degree) behaviour tracked consciousness, then we might be able to confirm or refute a report. But with no agreed-upon metric or theory of consciousness, we cannot do so. Irvine (2012) explains:

> With no other points of reference, even extreme biases cannot be discounted as bad, and introspective reports cannot be evaluated as more or less correct. That is, without knowing what the contents of consciousness actually are, and having no other methods with which to compare introspective methods, there is no clear way of establishing when introspective errors are made, or when subjects are 'correctly' reporting their experiences (p. 634).

Timmermans and Cleeremans (2015) make a similar point:

> … we do not know of any instrument or method that makes it possible to measure the contents of awareness directly (Seth et al. 2008). Having such an instrument (i.e. a consciousness-meter) would make it possible to establish clear relationships between an external state of affairs, people's subjective experience of this state of affairs, and their overt behavior.

> However, neither does such an instrument exist nor can we conceive of any way of building it ... (p. 21).

Essentially, because a) subjective reports are biased and not reliably indexed onto the presence of consciousness, and b) we do not have an independent measure with which to verify these reports, it is c) very difficult—if not impossible—to determine when (or if) a report has exhaustively captured all instances of degraded conscious information. Are blindsight subjects and DF *actually* unconscious of a stimulus when they say they do not experience it? We cannot precisely rule out—beyond a reasonable doubt—that all degraded instances of conscious information are reliably above the subjective threshold. Perhaps when a blindsight subject or DF reports 'no', they are conscious of the stimulus, but the conscious information is so degraded that it goes unreported.

Indeed, the upshot of what we have covered so far in Chapter Two is that we should not confuse an absence of evidence as evidence of absence (Timmermans & Cleeremans, 2015). My claim is not that blindsight subjects and DF *are* conscious of the stimuli, my claim is that it is an open question. Subjective measures are not reliable enough in distinguishing between unconscious processing and *degraded* conscious information for us to know either way.

So, according to our unconscious report card introduced in Chapter One, blindsight and DF record a question mark next to 'is unconscious':

	Blindsight	**DF**	**CFS**
Is unconscious	**?**	**?**	
Is sensory			
Involves constancy			
Is individual-level			

Before moving onto Chapter Three, though, where I analyse whether the alleged instances of unconscious perception satisfy Burge's criteria for perception, let us briefly turn our attention to whether CFS records a 'tick' on our report card.

2.2 Less Doubt for Continuous Flash Suppression

Firstly, instead of indicating *unconscious* perception, b-CFS studies—such as Yang et al.'s (2007) study—may only indicate a differential conscious detection threshold. For instance, similar to how a shiny object will catch your attention quicker than a dull one, perhaps a fearful face which "breaks" into consciousness quicker than a neutral one, merely informs us that we have a lower detection threshold for consciously perceiving fearful faces. Thus, there is a variable *detection* concern for whether b-CFS studies indicate unconscious perception rather than a variable *criterion* concern.[23]

Secondly, it *is* plausible that non-Mondrian images under the effects of "normal" CFS studies (non-b-CFS studies) are consciously detectable, but the images are so fleeting or degraded that they fail to meet the subject's criterion for a 'hit'. Stimuli might not completely achieve dominance over one's conscious vision and are thus 'missed' simply because the subject is not confident that they saw the faint signal. Indeed, perhaps CFS is not so binary that a subject is either fully conscious or fully unconscious of a stimulus. Conceivably, there is a grey zone between the suppression and dominance of a visual field where fleeting and degraded conscious information goes unreported (Yang et al., 2014). Therefore, *insofar* as CFS studies only use subjective measures to claim subjects are not conscious of a stimulus, it is a live possibility that they—just like blindsight subjects and DF— fail to report low and degraded conscious information. However, as flagged in Chapter One, some CFS studies use both subjective measures *and* objective measures to claim stimulus suppression. For instance, subjects in Jiang et al.'s (2006) study reported not seeing the stimulus and they performed at a chance level under a 2afc task (d' = 0). How should we interpret this?

We *could* press the point that it is far from clear whether performance in 2afc tracks consciousness. For example, even if CFS subjects displayed d' > 0 it would not prove conscious perception: objective measures are not reliably exclusive. Therefore, perhaps performing at chance levels of success in 2afc does not necessarily indicate unconscious perception either. It may instead merely indicate that the stimulus information is so degraded or fleeting that it cannot be used for above chance discrimination. When a signal is weak or

[23] For a longer discussion on b-CFS studies see the debate between Phillips and Block (2016). Unfortunately, I do not have the scope to rehash their debate in any detail; nevertheless, I take the outcome of it to be that it remains an *open* question whether b-CFS studies demonstrate unconscious perception. Also see Stein (2019) for a recent review of b-CFS.

minimally detectable, subjects may fail to use the degraded conscious information for successful 2afc task performance (see Yang et al., 2014, Lin & Murray, 2014, and Overgaard, 2017). Once again, perhaps an absence of evidence is not evidence of absence.

However, I think that trying to press the point Jiang et al.'s subjects might be conscious of the nude stimuli borders on the type of doubt that I have been trying to avoid. I am only interested in *reasonable* doubt, and I am not sure this qualifies. While it is possible d' = 0 does not infallibly tell us that a subject is unconscious of a stimulus, it does not suffer from the problem of the criterion the way reporting 'no' in subjective measures of consciousness does. Therefore, the problems faced by CFS in claiming subjects are not conscious of a stimulus are less severe than the problems facing blindsight and DF. Indeed, it seems plausible that not showing any objective perceptual sensitivity in an unbiased 2afc task—while not being infallible—*reliably* indicates that a subject is not conscious of a stimulus (see Irvine, 2013a, 2013b).[24] In other words, while there is certainly room for doubt, perhaps it is beyond a *reasonable* doubt that subjects in Jiang et al.'s study are unconscious of the nude stimuli.

In sum, while it is an open question whether blindsight subjects and DF satisfy 'is unconscious' on our unconscious report card, I propose that in regards to CFS, it is fair to give it a 'tick':

	Blindsight	**DF**	**CFS**
Is unconscious	**?**	**?**	✓
Is sensory			
Involves constancy			
Is individual-level			

Of course, as we are about to see in Chapter Three, if a subject's perceptual sensitivity is zero, then we lose grounds for thinking the subject is perceiving. Recording d' = 0 might increase our reason for thinking a subject is unconscious of a stimulus, but, by the same token, it *decreases* our reason for thinking they are perceiving.

[24] Keep in mind that even *if* d' > 0 is reliably exhaustive, it does not imply that perceptual sensitivity is also reliably exclusive. A measure can be exhaustive without being exclusive.

I now turn our attention to whether the alleged instances of unconscious perception satisfy Burge's criteria for perception. I do so by bracketing the concerns from this chapter, and asking: suppose each alleged instance *is* unconscious, are they even perceiving?

Chapter Three: Are the Alleged Instances of Unconscious Perception *actually* Perceiving?

Chapter Two demonstrated that we have reasonable doubt whether blindsight and DF are unconscious of the stimuli they seem to be perceiving. In Chapter Three, however, I move forward *as if* each alleged instance is unconscious. I assume for the sake of argument that they are unconscious and show that, regardless, they still vitiate Burge's criteria for perception.

With 'is sensory' and 'involves constancy' granted in Chapter One, failing or passing Burge's criteria boils down to whether they are attributable to the individual-level. Currently, then, our unconscious perception report card looks like this:

	Blindsight	DF	CFS
Is unconscious	?	?	✓
Is sensory	✓	✓	✓
Involves constancy	✓	✓	✓
Is individual-level			

In what follows, I argue that the alleged instances of unconscious perception are subpersonal processes and are thus *not* individual-level phenomena. I call this the 'substantial claim'. Nevertheless, I also aim to show that if one is not willing to commit to the substantial claim, then one must at the very least commit to the 'weaker claim'. The weaker claim holds that it is fuzzy or unclear whether the alleged instances are personal level or subpersonal level phenomena. The upshot, though, is that regardless of which claim we commit to, one is forced to concede that there is indeed reasonable doubt whether we have unconscious perception.

The substantial claim is derived from two points. One: that perception is a personal level phenomenon and, two, that the cases put forth as instances of unconscious perception are *not* personal level phenomena. This chapter will analyse each point in turn. First, I will explain what the personal and subpersonal distinction is and where perception fits into it. Second, I will explain what attribution to the personal level demands. What makes a

representation attributable to the individual? What criteria should we demand when claiming attribution?

I put forward an answer that claims the critical criterion is whether the content of a representation is available for intentional coordination. With that, we will see how the alleged instances of unconscious perception—instances such as blindsight, DF's visual form agnosia, and CFS—do not meet this criterion for attribution and thus fail to satisfy Burge's criteria for perception. First, though, what is the personal/subpersonal distinction?

3. 1 The Personal/Subpersonal Distinction

The personal level refers to whole persons as entities and what they do, whereas the subpersonal level refers to the subparts of a person and their physical processes (Dennett, 1969; Elton, 2003). In other words, the person is a type of integration of the parts into a coherent whole (see Bayne, 2013). However, the personal level is not just the sum of its parts; instead, similar to how a mass of gas stands to its component molecules, the person is a mental manifestation of its physical subparts (Hornsby, 2000).[25]

Personal level states of mind can include perceptual states like seeing, hearing or tasting, and it can also include cognitive states such as desires, beliefs, and memories (Shea, 2013). When we say, for instance, that John 'wants to go to Europe', we are referring to the person, John, as a mental entity with desires and dreams; we are not referring to a subpersonal physical part of what John's brain is doing. Moreover, the personal level can also include a person's emotional states and sensations (Shea, 2013). Dennett (1969) introduced the distinction with the example of pain:

> When we abandon mental process talk for physical process talk we cannot say that the mental process analysis of pain is wrong, for our alternative analysis cannot be an analysis of pain at all, but rather of something else - the motions of human bodies or the organisation of the nervous system. Indeed, the mental process analysis of pain is correct. Pains are feelings, felt by people, and they hurt (p. 90).

[25] Note that the personal/subpersonal refers to animals too. We could quite easily interchange 'personal level' with 'animal level' or 'subpersonal' with 'subanimal'.

When we refer to John's pain as a mental phenomenon, we are referring to activity at the personal level. When we refer to the physical processes underlying John's pain, we are referring to the subpersonal mechanics and motions that leave the person out of the picture.

Subpersonal level phenomena are the causal or mechanistic events in parts of the brain or nervous system. When John's primary visual cortex is receiving information or sending signals, the activity is a subpersonal event. John, as a person, does not operate the subpersonal machinery; he does not coordinate one subsystem to pass a message onto the other. Consider again John's thought about 'wanting to go to Europe'. The underlying physical, mechanistic subpersonal account of this could be something like: 'The visual cortex processed a stimulus which sent a signal to the amygdala that released a particular chemical which caused various neurons in the neocortex to fire.'[26] This subpersonal account does not make John, as a person, intelligible to us. For that, we need to switch levels and analyse his thought from the personal level. For instance, we could say something like: 'John saw an image of the Swiss Alps which gave him a peaceful feeling, so he thought he would like to go to Europe one day and experience the Alps in person'. In other words, if we are considering the physical mechanisms of a part of the person's brain or the events of a person's nervous system, we are not considering the person, but if we are considering John's thoughts and desires, we are considering personal level phenomena.

Drayson (2017) relates the personal/subpersonal distinction explicitly to perception by giving us a detailed example of how subpersonal mechanisms can help the person to perceive, without actually perceiving itself. She highlights the following subpersonal tasks which facilitate a person to perceive *depth*:

> The first task, finding matching points in the two retinal images, might require looking for patterns of light intensity that indicate the edges of objects. This task can be further decomposed into the sequence of steps required to calculate the points at which light intensity is changing most rapidly. Once the visual system has estimated which points in the retinal image correspond to the same part of the scene, it can perform the second task, calculating the disparity between them. The third task, recovering the three-dimensional scene, can also be decomposed into distinct tasks: figuring out the separation

[26] This is not designed to be anatomically correct; instead, it is just an example of a physical process that leaves the mental phenomena of the person out of the equation.

> distance between the two eyes, for example, and using gaze-angle to estimate fixation distance, then combining these figures with the disparity information to produce the three-dimensional description (p. 2).

In these tasks, the *person* is not figuring out correspondence points, calculating disparity, or recovering information; instead, the tasks are confined to the subpersonal visual system. Moreover, while not, *in themselves*, perceiving, these various subpersonal mechanisms explain the person's capacity to perceive. Consider, for example, the similarities with water, where water is akin to the person, and its elements (hydrogen and oxygen) akin to the subpersonal parts of the brain. Just because water is wet, it does not, therefore, follow that the hydrogen atoms or the oxygen atom are, in themselves, also wet (Drayson 2017). Instead, the wetness is a manifest property of its non-wet elements at the sub-water level. The same analogy can be given for a clock. The nuts and bolts which constitute the clock on, say, Big Ben in London, do not tell the time *per se*, but they enable the clock as a whole to tell the time (Bennett & Hacker, 2003). Likewise, the subpersonal parts of the brain which constitute the personal level do not perceive, but they allow the *person* to.

Hurley (1998) makes a similar point when she states that personal level perception can "... be seen as carried by such subpersonal processes, or vehicles of content. But the properties of subpersonal processes of vehicles of content, cannot simply be projected into personal level mental content or vice versa" (p. 3). And again:

> While some such processes or events might conceivably be identified with acts of the person in whose body the events occur, we have no reason to regard all processing events involved in perception as the acts of the person in whose body they occur. To do so is to confuse brain activity with intentional acts, information processing with agency … to conflate personal and subpersonal levels of description (Hurley, 1998, pp. 78-9).

Klotz and Neumann (1999) also state that:

> … perceiving is something that a person or an animal does, not something that can be properly ascribed to stages, sub-systems, brain areas, or the like. The triggering of a sneeze by an external stimulus does not imply that the reflex center that controls it "perceives" the stimulus (p. 976).

The point is that the personal/subpersonal distinction is significant because it gives us a naturalistic way of conceptualising, making intelligible and referring to persons that is non-physicalist, non-dualist, and non-supernatural (Elton, 2003). Instead of just brute causal and physical facts in the brain, it provides us with a framework for rationalising, justifying and understanding a person's actions in virtue of their mental states (Shea, 2013). Indeed, through the lens of the personal/subpersonal distinction, we can make sense of a person's intentional state of mind that may otherwise go missing. For example, Dennett (1969) asks us to imagine that we could create a complete "scientific story about synapses, electrical potentials and so forth that would explain, describe and predict all that goes on in the nervous system (pp. 87-8)", doing so, would mean that "all the motions ... of the animal caused by the activity of the nervous system would be explicable and predictable" (p. 88). However, even if we could achieve such a task, "such a story would say nothing about what the animal was doing" (p. 88). Namely, the story "would be mute on the topic of the actions (as opposed to motions), intentions, beliefs and desires of its subjects" (p. 88). To unmute the story, therefore, we need to switch from the physical mechanics of the subsystems to the mental *intentions* of the person.

Intentionality, in a philosopher's sense, generally refers to a thought or state of mind that is *about* a particular property or an object (Brentano, 1874; Searle, 1983). When John is thinking about Europe, for example, Europe is the intentional object of his thoughts. However, while the concept of intentionality is simple enough to grasp, the nature of it is disputed. As Davies (1995) points out, intentionality can come in many varieties of aboutness, and philosophers have taken different positions on which variety best captures the concept. Perhaps intentionality is a genuine and intrinsic aspect of the natural world that we can identify in many domains, all the way from persons to, say, thermostats. When John decides to go to Europe because he has a desire to do so, it is in virtue of this desire, and not anything else—such as a physical state of his brain—that he decides to go. In other words, it is in virtue of a *genuine* intentional state of mind that he chooses to go to Europe.

However, we might split the literal or genuine intentionality into multiple subcategories. Personal level intentionality might be a particular variety that is distinct from thermostat intentionality. Perhaps the personal level variety is capital I 'Intentionality', whereas when the thermostat has a belief *about* the temperature in the room, this is merely a lower case i

'intentionality'. Both are genuine and identifiable features of the natural world, but they are distinct from each other; or, at the very least, they are on opposite ends of the spectrum.

Alternatively, we can take the instrumentalist position that while speaking about intentionality may be useful for understanding aspects of the natural world (such as persons and their actions) it is just that: a useful fiction. It is a meaning, or pattern (see Dennett, 1987) that humans have procured to help us navigate the world. Similarly, for instance, to how we have created days of the week. The day Tuesday is not an intrinsic or genuine part of the natural world, but it is a useful fiction, one that we have attached meaning onto. Perhaps, as Elton (2003) speculates, the concept of intentionality is akin to the way we use days of the week.

Finally, we might follow Hornsby's (2000) position which claims that some varieties of intentionality are genuine, while others are merely metaphorical or *as if* intentionality. On this interpretation, the personal level has literal intentional states of mind that are intrinsic to the natural world, whereas the subpersonal or thermostat does not. Instead, it is merely useful to speak *as if* the subpersonal has intentionality. When the visual cortex communicates with the amygdala *about* a stimulus, the visual cortex and the amygdala do not genuinely have intentionality; instead, we merely talk *as if* they do. Hornsby (2000) explains:

> There is a real difference between adopting the intentional stance towards persons and adopting it towards other things. Adopting the stance towards persons is a matter of bringing a full range of common sense psychological and other concepts to bear on beings to which it is suited; adopting the stance towards other things—towards subsystems, thermostats or whatever—is a matter of treating them *as if* they had some of the intentional properties that persons have (p. 20) [Hornsby's emphasis].

Whether we ultimately agree with Hornsby's position or not, it does seem that she has elicited a valuable insight. Namely, there is a difference between adopting the intentional stance towards persons and adopting it towards subpersonal parts or thermostats. The way we refer to intentionality can and should come in different degrees and varieties *regardless* of whether we think intentionality is real, instrumentalist or a combination of both. For our purposes, this distinction is the important lesson to take from intentionality. Nothing I go onto say below in section 3.2 regarding the criterion for personal level attribution hinges on which position we settle on. What does matter, though, is that we acknowledge that personal level

intentionality comes in a different variety or degree of intentionality to that which is not at the personal level. Any explanation of how to attribute perception to the personal level *must* capture this difference. Perception is a mental phenomenon that is a personal level intentional state towards or about an object or property. In 3.2 I offer a criterion for attribution to the personal level that captures this distinction in intentionality while at the same time shows how the alleged instances of unconscious perception are not personal level intentional.

Furthermore, in accepting that personal level intentionality is distinct from other types of intentionality—while not committing ourselves to whether intentionality is real, instrumentalist or *as if*—we can preserve the distinction between personal level mental phenomena and the subpersonal physical system. We do so while also not hinging the criterion for personal level attribution on a controversial or unsubstantiated conception of intentionality.

3.1.2 The Upshot

So far, we have circumscribed what the personal/subpersonal distinction is and why it has ramifications for perception. Perception is not just a subpersonal physical process in the brain; instead, it underpins how a *person* navigates their environment. As Burge (2010) states, “if perception is to ground this explanatory role [of action], it must be attributable to the individual” (p. 371). Indeed, both sides of the unconscious perception debate—specifically Phillips and Block—agree with Burge that perception is a personal level phenomenon (Phillips & Block, 2016). The disagreement, however, begins at how or when to attribute an alleged instance of unconscious perception to the personal level. What is the criterion for personal level attribution? When can we attribute a mental state to the person? It is the answers to these very questions which Phillips disagrees with Block (and Burge) on.

The disagreement is substantial because where one stands on unconscious perception could boil down to how one interprets the criterion for personal level attribution. Block (2016) and Burge (2010), for example, understand the criterion in such a way that they conclude the alleged instances of unconscious perception *are* personal level phenomena; therefore, the alleged instances *are* perception. Phillips (2016, 2018), on the other hand, understands the alleged instances *not* to be personal level; therefore, they are *not* instances of perception. Can we make a metric for attribution that helps us move forward in the debate? Why does

Phillips disagree with Block and Burge over the criterion for attribution, and how can we resolve this problem?

Below in 3.2, I present what I think is the most credible criterion for attribution to the personal level. I will show that if we accept the premise that perception is a personal level phenomenon, then the perceptual content *must* be available for intentionally coordinated action. Note that when I say 'intentionally coordinated' I am now using 'intentional' in the common usage sense (of intending to do something). As we will see, though, this common understanding will importantly capture the demands of the philosopher's personal level intentionality canvassed above.

If my criterion for personal level attribution is accepted, then I will also show that my substantial claim that the alleged instances of unconscious perception all fall short of personal level attribution should also be accepted, and therefore they do not satisfy Burge's criteria for perception.

3.2 Attribution to the Personal Level

In attempting to answer the question 'what is it for a perceptual state to be personal level?', both sides of the debate have taken the strategy of identifying a specific functional role that all and only personal level perception is capable of satisfying (Berger & Mylopoulos, 2019). This is also my approach.

Firstly, Burge's (2010) answer to what makes a perceptual state attributable to the personal level is: "where a sensory state, non-perceptual or perceptual, can *initiate action* by an individual, it is attributable to the individual" (p. 373) [my emphasis]. Likewise, by referencing DF's visual form agnosia, Block (2017) takes the same stance:[27]

> In the case of DF there is as definitive an answer to the personal/sub-personal question as one is going to get for this question: her unconscious visual representations are *"her*" representations because they guide *her* actions (p. 8) [my emphasis].

[27] As briefly mentioned in Chapter One, Block also suggests that when a sensory registration reflects personal level values we gain reason to attribute the registration to the personal level. I address this below in 3.2.1.1.

Under Burge and Block's conception, the "definitive" criterion for attribution to the personal level is whether the representation guides (or "initiates") the individual's ("her") actions. Moreover, Phillips (2018) states that "availability to central agency comprises our best evidence for attributing a representation to an individual" (p. 494). Perhaps it is the case that both conceptions are consistent with availability to a central agency, but disagreement stems from how to understand this availability. Therefore, I aim to provide clarity by arguing that the availability of the content of a representation for intentional coordination is the key to attribution.

My move will be twofold. First, I will split guidance into two distinct subtypes. Burge and Block's above quotes indicate that there is only one type of guidance which is relevant for attribution to the personal level. In what follows, I show there are *two*. The first type of guidance is robotic, automated and located at the subpersonal level; these are only mere happenings in the person's subsystems that can indeed affect the personal level, but occur without personal level intentionality. The second type of guidance, however, is the opposite: the *person* intentionally coordinates their movements. Contrary to Burge and Block, rather than merely guiding a person's actions, the guidance must be of the right subtype for attribution to the personal level.

Let us say, though, that Burge and Block are persuaded by my criterion for personal level attribution and yet do not see it as a problem for the alleged instances of unconscious perception. If so, then my second move is to show that if one does agree with my criterion, then the alleged instances are *not* personal level phenomena. To recap from Chapter One, the argument formulates as follows:

1. Perception is a personal level phenomenon.
2. Attribution of a personal level state requires the *availability* of the content of a representation for intentional coordination.
3. The content of the representations from the alleged instances of unconscious perception is not available for intentional coordination.

Therefore:

4. The alleged instances of unconscious perception are not instances of perception.

I take this argument to be valid without further explanation. What does require explanation is why this argument is sound. Premise One, as already mentioned, is accepted by both sides

of the unconscious perception debate, so I will not spend more time defending it here. Premises Two and Three are where the disagreements lie. If someone is to dispute my conclusion, then they need to find fault with one of these premises. Thus, the remainder of this chapter will be spent defending them.

I use 'representation' to denote a psychological information-bearing state, and 'content' to refer to the objects or properties that the representation is about. So, essentially, my argument is claiming that a representation is personal level perception *if and only if* its contents are available for intentional coordination. Defending Premise Two begins by expounding what the concept of attribution, properly understood, demands. This task will be significant because once we understand these demands, we will be in a better position to see how and why intentional coordination follows as the necessary and sufficient criterion for personal level attribution.

3.2.1 Premise Two: Attribution of a personal level state requires the *availability* of the content of a representation for intentional coordination.

The concept of attribution is closely related to the concepts of responsibility and control. If I am in control, I am an apt candidate for responsibility, and by the same token, if I am responsible, then I am an apt candidate for attribution. To develop this connection further, consider two examples from Fischer and Ravizza (1998).[28]

Imagine that you are the captain of a ship and are intentionally guiding the ship towards a harbour. You recognise, for example, the reasons for the ship needing to dock at the harbour instead of, say, a rocky shoreline. With these reasons in mind, you then intentionally coordinate your actions by directing the ship accordingly. Therefore, because you (the person) own the cognitive power used in this guidance control, we can hold you responsible for the direction of the ship (Fischer & Ravizza, 1998). Likewise, because you are in control of the direction and thus responsible, we can also *attribute* the direction of the ship to your guidance. However, imagine if all of a sudden the steering wheel stops working, the rudder breaks or the winds become so intense that you are no longer in control of the direction of

[28] While I think that when we have a case of responsibility we always have a case of attribution, I remain neutral as to whether the reverse holds true. My model goes like this: Control → Responsibility → Attribution. Also, note that Fischer & Ravizza (1998) develop a theory for *moral* responsibility; however, I am only leaning on them to develop a connection between responsibility and attribution while remaining neutral on any claims about moral responsibility. Below in 3.2.1.1, I address this distinction in more detail.

the ship. In such a case, you are no longer responsible for the direction either; your cognitive control is bypassed, and the direction occurs independently of your influence. Therefore, with the captain's control and responsibility of the ship's direction dissolved, attribution is also dissolved. As Fischer and Ravizza (1998) explain: "he [the captain] does not guide the boat, the boat's movements are no reflection on him; rather, they are entirely attributable to the winds" (p. 221). This example gives us a good indication of why guidance should be split into two distinct subtypes and how these subtypes are relevant to the personal/subpersonal distinction.

On the one hand, the captain directs the ship according to his guidance control. On the other hand, though, the winds' motions move the ship independently of the captain's guidance and are thus being moved in such a way that is not attributable to him. Similarly, if a person (just like the captain) controls their guidance according to the content of their representations, then we can attribute that guidance to him. Whereas if subpersonal mechanisms move a person independently of their control (just like the winds), the person is thus bypassed and moved in such a way that we cannot attribute the movement to them. The critical criterion for attribution to the personal level is that the person must have the capacity to be in control of the guidance. Another example from Fischer & Ravizza (1998) builds on this point.

Imagine that an evil scientist has implanted a chip in your brain without you knowing. The chip is so clever that it forces you to act according to the scientist's wishes; if the scientist wanted you to commit murder, you would do so. Via the chip, the scientist has taken over the exploitation of the visual representations to guide your actions. Consequently, if you were to murder someone because of the chip's guidance, it would not be appropriate to hold you responsible—you lack the control needed for responsibility—likewise, because you lack control and responsibility, we would also not attribute the murder to you either. Guidance is not, in itself, a reason for attributing a representation to the individual; instead, the guidance must be available for control *by the person* before we can attribute it to them.

So, the concept of attribution, properly understood, demands control. If we have control and responsibility, *then* we also have attribution to the person. This control, moreover, requires that the person cannot be bypassed; the use of the content cannot occur independently of them. If we accept this connection, then it follows, I submit, that we have attribution to the person if and only if the content of the representations is available for intentional coordination. Indeed, the difference between intentionally coordinated guidance and

unintentional guidance is the fundamental cognitive control an intentional action has. The captain of the out of control ship and the person with the implanted scientist's chip, for instance, did not have the control required for attribution because intentional coordination was not available to them. Likewise, consider the difference between intentionally making a coffee and pupil dilation.

When I grind the coffee beans, steam the milk, and pour the coffee into a mug, these are intentionally coordinated actions directed towards a goal. I am in a position to use the content of my representations in a variety of ways. I can control, for example, how much milk I want in my coffee, I can decide to let the coffee cool down before drinking or I can give the coffee to someone else to drink. The content is available to me in such a way that I can intentionally coordinate my actions. However, if I walk into a room, and the light is slightly dimmer than it was outside, my eyes will often automatically dilate independently of my guidance. The information needed for dilation may bypass the personal level and only be available to my eyes. With that being the case, though, then pupil dilation is something that happens to me or within me, rather than being something that *I* do. As Frankfurt (1978) explains:

> The dilation of the pupils ... the occurrence of this movement does not mark the performance of an action by the person; his pupils dilate, but he does not dilate them. This is because the course of the movement is not under *his* guidance. The guidance in this case is attributable only to the operation of some mechanism with which he cannot be identified (p. 159) [Frankfurt's emphasis].

In other words, the reason why pupil dilation is not attributable to the personal level is that it is not available for intentional coordination. Whereas, the reason why making the coffee *is* attributable to the personal level is because the content is available for intentional coordination (also see Drayson, 2014, for a similar line of thought). To reinforce this point, consider two more examples.

Imagine someone in an epileptic seizure where a sudden burst of electrical activity affects the brain, leading to (amongst other consequences) the person exhibiting physical convulsions and abnormal behaviour. In these instances, we do not attribute the behaviour to the person *as such*; it is clear that they have lost the guidance control necessary for

attribution. Again, I suggest the reason for this is because they cannot intentionally coordinate their actions. Instead, subpersonal physiological responses are taking over their behaviour. However, once (or if) intentional coordination becomes available to the person, then we also, in the same instance, regain both the necessary and sufficient reason to attribute the action guidance to them.

Moreover, imagine that you have recently become a vegetarian. Up until this change, though, your favourite food was bacon. One morning shortly after this change of diet, you are in a café and smell bacon frying right next to you; the aromas trigger a physical response that is so strong your hand starts to reach out and grab a piece reflexively. However, then you remember you do not eat meat anymore, so you begin to intentionally coordinate your behaviour away from eating the bacon. Insofar as the content from the representations is *available* for intentional coordination, then you have the necessary cognitive control for personal level responsibility and attribution. However, if you lacked this ability, say you were a mindless automated drone that merely acted according to sensory-stimulus driven reflexes or reaction, then we would not, I think, attribute the guidance to the personal level. Instead, we would recognise that the movement is a result of subpersonal components and their mechanical responses.

Finally, before I add a few clarifications and address some objections, consider the connection between intentional coordination and cognitive integration at the personal level. When we can cognitively integrate various ideas, thoughts, beliefs and propositions into one coherent representation, it seems to align well with the personal level (Bayne, 2013). For instance, if I arrange to meet a friend at a bar in Melbourne at 5pm on Friday, various cognitive processes have to integrate for me to understand what meeting at the bar entails. I have to know, for example, Melbourne is accessible for my friend and I to get to, and that neither of us will be in, say, Brisbane at the time. I also have to know that showing up on Friday means not showing up on Thursday and that 5pm means not arriving earlier or later in the day. Perhaps I am also aware that meeting at the bar will mean I need to spend money to be a patron there, so I will know that by agreeing to meet I have the cash available to do so. The list could (and does) go on, but the point is this: if the personal level involves a kind of cognitive integration, then perhaps the higher the capacity we have for intentional coordination, the higher the level of cognitive integration. That is, the *more* we can intentionally coordinate the content of our representations, the *more* our cognitive economy becomes integrated. I picture the connection like a coat hanger, where the coat hanger holds

or props up the clothes from falling; likewise, the availability of intentional coordination props up cognitive integration from falling into isolated compartments. The stronger or more sturdy the coat hanger is (intentional coordination), the stronger and more sturdy the clothes stay up (our cognitive economy integrates).

I will say more about cognitive integration in Premise Three. Meanwhile, I have set out the main structure of my claim in Premise Two, namely: intentional coordination is the key to achieving the control required for attribution to the personal level. Now, I add various clarifications, caveats and address some objections.

3.2.1.1 Clarifications, Caveats, and Objections

Firstly, could intentional coordination be a sufficient criterion for the personal level without being a necessary one? That is, do cases such as locked-in syndrome show that we can attribute mental states to the individual regardless of whether they can intentionally coordinate? I do not think so. The content of the locked-in syndrome patient's representations is still *available* to them for intentional coordination; the only difference between them and an able-bodied person is that the pathway from the brain to the limbs is inhibited. The content itself is still suitably placed for intentional coordination, thus satisfying the requirement for attribution to the personal level. Indeed, above in 3.2.1, I primarily concentrated on how intentional coordination ensures that we have the control and responsibility required for attribution. Here, though, I emphasise and clarify that if this is the case, insofar as the content of our representations is *available* for intentional coordination, we have personal level attribution.

Consider, for example, a fragile glass that has a disposition for shattering if dropped. As Armstrong (1981) points out, this disposition to shatter is an intrinsic feature of the glass. Even if we put layers of padding around it so that it will no longer shatter when dropped, nothing actually changes about the glass; its natural disposition to shatter remains the same. We can say something similar for the locked-in patient. That is, their intrinsic state is still attributable to the personal level insofar as the content from the representations is *available* for intentional coordination. The only reason they might not be able to intentionally coordinate as an able-bodied person might is due to *non*-intrinsic factors which—just like the padding around the glass—does not change anything intrinsic about their mental state. As I explain with further detail below in Premise Three, this is different from someone like a

blindsight subject or DF, who is intrinsically not disposed to intentionally coordinate. The content for intentional coordination is *not* available to DF the way it is for the locked-in patient.[29]

My next point is to clarify the connection between attribution and responsibility. Even though I think the connection holds, I do not mean to say that attribution entails *moral* responsibility. Consider, for instance, someone with a debilitating drug addiction. Such an addiction might mean that the subpersonal cravings are so intense that the addict is, in a sense, compelled to obtain and take drugs regardless of any personal level value or preference for quitting. Does this mean that we are forced into the untenable position of not attributing perception to the personal level simply because the addict has (perhaps) lost the control necessary for *moral* responsibility? No. The addict can still intentionally coordinate the content of his representations. The personal level is still responsible for his guidance while not necessarily being morally responsible. As Schroeter (2004) explains:

> When the execution of his action is about to start and throughout its unfolding, the addict is *aware* of what he is doing or what he is about to do: he entertains an *action plan*—that is, a representation of the task to be performed by his action (p. 640) [Schroeter's emphasis].

Indeed, the content of the representations do not bypass the addict's personal level; he must still purposefully guide, monitor and coordinate the unfolding of his actions to secure the realisation of his goal. For instance, during the process of obtaining the drug, he might need to complete a difficult reverse park (Schroeter, 2004). Also, when, say, carefully guiding the syringe to hit the targeted vein, he must still intentionally coordinate his movements; otherwise, the action would stop or be derailed. In one sense, then, the addict is so compelled to complete his task that he is not *morally* responsible for doing so. In another sense, though—the sense I am trying to elucidate—the representations used to complete his tasks and to realise his goal, is intentionally coordinated; thus, the personal level is responsible for completing the action. We can attribute control, responsibility and attribution for an action to the personal level without also needing to attribute moral responsibility.

[29] I say more about this in Chapter Four, where I show how we can still detect intentional coordination in locked-in patients.

We can say something similar for various animals. When a cheetah tracks a gazelle, for example, the cheetah is not morally responsible for killing it. Instead, when it purposefully tracks its prey and uses content from its representations to assess a variety of options (such as when and how to obtain its prey), the personal level is responsible for executing the action. The cheetah is not bypassed, and has the required intentionally coordinated control to attribute personal level perception, without needing to also attribute *moral* responsibility.

With that said, let us move onto another issue, and one that was flagged in Chapter One, namely: is a sensory response which aligns with personal level values and understanding sufficient for attribution? We saw that Block certainly thinks this is plausible. To recap, he says (in Phillips and Block, 2016): "when a sensory registration reflects personal-level *understanding* ... or *values* (Jiang et al., 2006), that is a reason to think it is a personal-level perception" (p. 23) [Block's emphasis]. However, while I am willing to concede that personal level values and preferences may not be irrelevant, I do not think they are necessary or sufficient.

Firstly, as clarified in section 3.1.2, one of the primary reasons why perception is a personal level phenomenon in the first place is that it underpins the *person's* ability to navigate their environment; it is the interface between external stimuli and a person using these stimuli for action. Therefore, it is reasonable that attribution of perception to the personal level should hinge on its role (of guidance) instead of its content.

Secondly, even though values may often overlap with the personal level—so at times can be a useful marker for identifying attribution—they are only useful insofar as when we intentionally coordinate we often do so in accordance with our values. Nevertheless, in regards to the addict, for instance, it does not seem that his "sensory registrations" reflect his personal level values and understanding, and yet he is unquestionably perceiving. On the one hand, the sensory registration of a drug aligns with his strongest desire, but merely aligning with a desire could indicate nothing more than a sensory-driven response and not, that is, perception by the individual. A moth, for example, might have the desire to go towards a light, but the sensory registration of the light does not indicate that representational content has integrated with a personal level. On the other hand, when the addict has a sensory registration that there is a drug in front of him which he feels compelled to take, this registration does not reflect his personal level value of, say, overcoming his addiction, and neither his understanding that he should not take the drug.

Therefore, I do not think the correct thing to say is that the distinction between the personal and subpersonal levels hinges on whether a sensory registration reflects or does not reflect our values and understanding. Values can be a useful indicator at times, but they do not seem to be the criterion which makes *the* difference between a mere sensory registration and perception which can be attributed to the personal level. If it did, then it would mean that content could bypass the personal level's guidance; it could be used independently of the person's coordination, and yet still be attributable to the personal level. If we take the connection between control, responsibility and attribution seriously, though, then this is implausible.

Finally, before moving onto Premise Three, it is worth briefly addressing Block's argument—which we first encountered in Chapter Two—that phenomenal consciousness can exceed or *overflow* cognitive access. Indeed, an overflow of consciousness potentially forms an objection to Premise Two. For instance, if a) as widely assumed, consciousness is only found at the personal level (see Elton, 2000, for an example), and b) there is an overflow of consciousness that we cannot cognitively access, then c) perhaps there is personal level perceptual content not available for intentional coordination. In other words, just as consciousness overflows cognitive access, so too, the personal level may overflow intentional coordination.

In response, note that it is far from clear—as was also implied in Chapter Two—that phenomenal consciousness overflows cognitive access. Kouider, Dehaene, Jobert, and Le Bihan (2007) explain:

> Given the lack of scientific criterion, at this stage at least, for defining conscious processing without reportability, the dissociation between access and phenomenal consciousness remains largely speculative and even possibly immune to scientific investigation (p. 2028).

Accordingly, it may be the case that there is no actual problem here for my claim which I need to respond to. A reasonable response to the objection from overflow is to simply insist that the burden of proof remains with the objector (also see Cohen and Dennett, 2011, Brown, 2014, Dehaene, 2014, Carruthers, 2017, for more on why claiming that consciousness overflows cognitive access is problematic). However, *if* there is an overflow

of phenomenal consciousness, then it is unclear why we should (by default) grant this overflow as an instance of perception. For example, if the overflowed consciousness is not available for intentional coordination, then why is it automatically attributable to the individual?

Perhaps the upshot of overflow consciousness—insofar as it is present—is that it merely informs us that consciousness is not always located at the personal level. Nevertheless, I *do* realise that this upshot is a bold suggestion and that in all likelihood not everyone will buy into it. However, note that it will only ever be in rare cases—if at all—where conscious content is not available for intentional coordination, i.e. not at the personal level. In fact, if there is no overflow of consciousness, then perhaps there are no *actual* cases where consciousness dissociates from the personal level. Either way, I think our default position should still be that intentional coordination marks the personal level, and if the consequence of this position is that consciousness is not *necessarily* personal level, then perhaps that is the bullet we must bite. Of course, no doubt there is much more which could be unpacked here; doing so, though, is beyond my scope. Instead, I offer the outline of two responses for further development. One, the burden of proof remains with the objector to show that phenomenal consciousness can overflow cognitive access. Two, perhaps the consequence of overflow consciousness (insofar as it exists) is that in very rare instances consciousness is not located at the personal level.

So, to briefly sum up Premise Two. I have split guidance into two distinct subtypes and only one type is necessary *and* sufficient for personal level attribution. A substantial amount of this claim rests on the relationship between attribution, responsibility and control. I have sought to show that if we take this relationship seriously, that is, if we pay due diligence to the demands of attribution *as a concept*, then intentional coordination is the key criterion for personal level attribution.

We are now equipped to apply my argument from Premise Two onto the alleged instances of unconscious perception. In doing so, I will further sharpen the force and understanding of intentional coordination by analysing how the alleged instances fall short of meeting its demands. Ultimately, I hope to show that if we accept intentional coordination as the necessary and sufficient criterion for personal level attribution, then we must commit to the substantial claim and reject the alleged instances as satisfying Burge's criteria for perception.

3.2.2 <u>Premise 3:</u> The content of the representations from the alleged instances of unconscious perception is not available for intentional coordination.

To defend Premise Three, I will measure the alleged instances of unconscious perception against a series of interconnected *indicators* for intentional coordination. The indicators are as follows:

(i) Integration
(ii) Flexible behaviour
(iii) Goal-directed behaviour
(iv) No contradiction or bypassing; the power of veto
(v) Self-prompting, including the ability to act spontaneously and on cue

There are a few qualifiers before we begin. First, I am not presenting this list as the final word on what indicates intentional coordination; I am merely putting it forward as a starting point and suggest that it *mostly* captures what we are looking for. The list will suffice for our current purposes and I am open to further refinement and revision later. Second, the degree with which each indicator is satisfied and the quantity of indicators satisfied serves to enhance or weaken the case for personal level perception. If an alleged instance satisfies these indicators, then I suggest that they can intentionally coordinate, and we can thus attribute personal level perception to them; if they do not satisfy them, though, then we lose grounds for attribution. Third, I will not always analyse each alleged instance against each indicator in equal detail, but, instead, sometimes take my analysis of one instance and assume it as a metric for the others. Nevertheless, I aim to do enough to demonstrate that we should plausibly accept Premise Three. Finally, and perhaps most importantly, I reiterate a point made at the start of the chapter; that is, the following analysis is given *as if* the alleged instances are unconscious. I assume for the sake of argument they are unconscious and ask: even if they are, can they intentionally coordinate?

(i) Integration:

Consider DF's ability to successfully insert mail into a slot (see Goodale and Milner, 2004, and Chapter One). When she does so, she merely guesses what the visual stimulus could be, has to be prompted to undertake the task, and is surprised when her actions are

successful. However, if the content of the representations were integrated into her cognitive economy, then I suggest she would have been able to use both the content from the mail and the slot to combine it with her working memory (which understands that mail goes into a mail slot) and complete the action without prompting or surprise. This ability, however, does not seem to be available to her.

Blindsight subjects do not fare any better. From their first-person perspective, they were merely guessing at what was presented to them (again, see Chapter One). They repeatedly denied seeing any stimulus in their blind field and had to be coaxed into using the content of their representations. Subject GY, in a Kentridge, Heywood, and Weiskrantz (1999) study, even accused his experimenters of not displaying any stimulus. As Kentridge (2011) later reported: “He steadfastly denied any knowledge of targets and, indeed, suggested at one point that there were no targets and that we were running some control condition” (p. 239). This reaction from GY indicates, I think, that the representation does not adequately integrate into the rest of his cognitive economy. Instead, the content of the representation seems to be confined to subcortical pathways (see, for instance, Weiskrantz, 1997; Cowey, 2010).

CFS subjects also fail to indicate integration beyond a reasonable doubt. At first glance, it might appear that when a fearful face “breaks” through suppression to consciousness quicker than a happy face, it reveals an integration of content with contextual understanding. However, as mentioned in Chapter Two, a more plausible explanation for b-CFS studies might simply be that the conscious detectability threshold for fearful faces is lower than happy faces. Therefore, a stimulus “breaking” through to consciousness quicker than another may only indicate differential timing in *conscious* integration. Moreover, as we will see below by analysing the further indicators of intentional coordination, it seems that subjects under the effects of CFS in Jiang et al.’s (2006) study do not indicate perceptual integration past the subpersonal circuits of the brain (such as the retina, amygdala and optic nerve).

(ii) Flexible behaviour:

Being able to intentionally coordinate involves the ability to draw on various behavioural capacities *flexibly*, appropriately, spontaneously and on cue (see Bayne, 2013). As Pennartz, Farisco, and Evers (2019) explain, flexible behaviour involves “building a

sensory-specific model of stimulus-action-outcome relationships, on prospective activity that anticipates future events, and the ability to make decisions on the fly, spontaneously, based on generalizable internal models" (p. 7). However, the alleged instances of unconscious perception do not seem to have this flexibility.

Firstly, subjects in Jiang et al.'s (2006) study could not successfully complete a 2afc task; they exhibited *no* perceptual sensitivity towards the nude stimuli (d' = 0). Therefore, the content of their representations was not available to complete the most simple of behavioural tasks, let alone to be used flexibly. Instead of the content opening an array of behavioural options to them, there were not *any* behavioural options available.

Secondly, while blindsight subjects can successfully complete 2afc tasks, they appear to be *mostly* restricted to these tasks, and so do not indicate a range of flexible behaviour. As we saw in Chapter One, and as Azzopardi and Cowey (1997) put it: "Blindsight is the rare and paradoxical ability of some human subjects ... to discriminate unseen stimuli in their clinically blind field defects *when forced-choice procedures are used*" (p. 14190) [my emphasis]. Indeed, if there is a cup of coffee in the blindsight subject's blind field, even though he might be able to tell us which spatial interval the cup is situated in (successful forced-choice), he cannot use the content of the cup for flexible behaviour. He would not, for instance, use the content to pick the cup up, get a teaspoon of sugar to mix in, gently take a sip, or pour milk into it, *not* because he lacks the motor ability (like a locked-in patient), but because the content is not available for him to do so.

Thirdly, even though DF might show more flexible behaviour than a blindsight or CFS subject, she still does not show enough to satisfy indicator (ii). Consider, for example, Clark's (2007) observations into the sorts of abilities which DF *lacks*. He says: "Suppose DF were to visually inspect a room, she would not ... be able to see that the ... object *over there* would fit into the ... space *over here*" (p. 586) [Clark's emphasis]. Such judgements, though, are the paradigmatic type of judgements we would expect someone to make if they can flexibly use the content of their representations. As Clark (2007) goes onto say:

> … we cannot ask DF a question about what might fit where ... But if we could somehow set things up so that she had a goal of, say, blocking up the space that is (as it happens) between the sofa and the table, my strong bet is that she would nonetheless fail to experience the visual scene in such a way as to

> make the potential fit visible to herself. She would then not, for example, proceed to move the furniture even on a 'hunch' that the goal would be achieved (p. 586).[30]

In other words, the content of the representations for subjects in the alleged instances fails to be available to them in such a way that it can be used for an array of flexible behavioural activities; thus, vitiating indicator (ii).

(iii) Goal-directed behaviour:

To once again draw on Pennartz et al. (2019), subjects "display goal-directed behaviour [as opposed to habitual behaviour] when this behaviour is driven by a representation of the expected consequences of action and depends on knowledge of actions being causal for obtaining a desirable outcome" (p. 7). So, because the CFS subjects in Jiang et al.'s (2006) study did not show *any* behavioural sensitivity towards the stimuli, it follows that they cannot use the representation to display any goal-directed behaviour either. If the content were available for goal-directed behaviour, then, at a bare minimum, we would expect subjects to have the capacity to use the content for driving their behaviour towards successfully completing the 2afc task. However, even though blindsight subjects and DF show behavioural sensitivity, they also do not exhibit goal-directed behaviour.

Think again of DF inserting mail into the slot. Her surprise upon completing the task indicates that she was unable to monitor her actions and know, for instance, that as she adjusts her hand to the appropriate angle, that this adjustment is apt for obtaining the desired outcome. Moreover, blindsight subjects' surprise at being told their guesses were correct also indicate that their behaviour was not goal-directed. As Dennett (2001) points out, we should take the presence of surprise quite seriously: "Surprise is a wonderful, dependent variable, and should be used more often in experiments; it is easy to measure and is a telling betrayal of the subject's *having expected something else*" (p. 982) [Dennett's emphasis]. When a subject is surprised that their action was apt for success, it reveals to us that the representation was not being driven in such a way that they could form beliefs about

[30] Note that Clark (2007) has a couple of aims when discussing DF. First, he is arguing that she lacks "a certain kind of link to personal agency" (p. 587). Second, that personal agency is tied to consciousness. I am strictly and only leaning on his argument in regards to the first aim.

their behaviour being “causal for obtaining a desirable outcome”, i.e. their behaviour was not goal-directed.

(iv) No contradiction or bypassing; the power of veto:

Consider studies from Pisella et al. (2000) and Danckert and Rossetti (2005) on the contradiction between automated guidance and intentional commands in the dorsal stream. Firstly, as we already saw in Chapter One (and see again Goodale and Milner, 2004), DF’s residual visuo-motor abilities rely on the dorsal stream. However, as Danckert and Rossetti (2005) explain, instances of *action*-blindsight—where blindsight subjects undertake visually guided movements or direct behaviour towards a blind field stimulus (such as pointing)—also rely on the dorsal stream. Therefore, the below studies have ramifications for our analysis of DF’s visual form agnosia *and* blindsight.

Pisella et al. (2000) showed that automatic corrections which resist intentional control might be the most specific feature of the dorsal stream. In their study, subjects were asked to point at moving visual targets which could unexpectedly change directions. However, when subjects tried to *not* keep pointing at the target when the direction changed, they would nevertheless automatically adjust their pointing according to the perturbations. This adjustment, therefore, contradicted their commands and indicated that the dorsal stream can operate independently of intentional control. Despite subjects’ best efforts to resist pointing, they still did so. Danckert and Rosetti (2005) further explain:

> Corrective movements were performed in the direction of the target jump in spite of the instruction to halt their movements. These corrections were considered to be automatic because they were produced spontaneously by naive subjects *against their own intention* to stop their movement in accordance with instructions. After touching the displaced target, subjects were fully aware of their mistakes and impulsively expressed frustration at their ‘error’. This ‘automatic pilot’ systematically activated during movement execution led subjects to produce ‘disallowed’ corrective movements (p. 1043) [my emphasis].

In other words, these studies demonstrate that the dorsal stream *can* act independently of our intentional coordination by contradicting and bypassing the personal level’s guidance.

Berger and Mylopoulos (2019), however, criticise the use of Pisella et al. (2000) and Danckert and Rosetti's (2005) studies to draw such conclusions. They do so on the grounds that the studies show "only that movements with a time course of under 300 ms are encapsulated from intention" (p. 24). Instead of giving a full rebuttal, I simply reiterate that automatically using representations independently of personal level guidance still seems to be a specific feature of the dorsal stream. Therefore, *insofar* as blindsight and DF's visual form agnosia rely on the dorsal stream for action guidance, and *insofar* as the guidance in the dorsal stream can act independently and contradictory to personal level intentional coordination, indicator (iv) is vitiated. Nevertheless, even if we were to disregard the studies from Pisella et al. (2000) and Danckert and Rosetti (2005), I think it still remains evident that the representations used by the dorsal stream can bypass or contradict guidance at the personal level. Two more examples show why.

Consider an analogy by Goodale and Milner (2004) regarding the similarities between a Mars rover and the dorsal stream (also see Clark 2007, 2009). Say you are in a mission control centre on Earth directing a rover on Mars. Your job is to review images sent to you from the rover and then command it to move towards items of particular interest in the distance, such as a rock. Once you send this command, however, the rover does the rest; its locomotion and computations kick in, and it moves towards the rock without you needing to direct its specific movements. Indeed, when the rover is moving towards the rock, it operates in a similar way to the dorsal stream i.e. independently of central command. Goodale and Milner (2004) further explain:

> The human operator doesn't have to worry about the real metrics of the workspace or the timing of the movements made by the robot [dorsal stream]; instead, the human operator has the job of identifying a goal and specifying an action towards that goal in general terms. Once this information is communicated to the semi-autonomous robot, the robot can use its on-board range-finders and other sensing devices to work out the required movements (p. 99).

In other words, just like the human operator and the rover, the personal level is not always privy to the specific content of the representations used by the dorsal stream in carrying out its computations; the information can bypass the person. Furthermore, what if the person

would choose a different track towards the rock *if* he could intentionally coordinate the rover's movements? Insofar as the rover operates similarly to the dorsal stream in *DF's* spared visuo-motor abilities, then such an option, I submit, is not available. For example, consider DF on a hike.

When DF is hiking, her dorsal stream automatically processes information that allows her to avoid stepping on an uneven or spiky rock (see Chapter One again). However, if the content of the rock were available to DF for intentional coordination, then she would have had the option of stepping on the rock *if she so chose*. It may seem unlikely that she would want to do so purposefully, but the point is that *if* she wanted to step on the rock and injure herself, then by automatically stepping around the rock, the action has contradicted what she would do if intentional coordination were available. Hence, the guidance is of the wrong subtype. The same can be said about CFS subjects.

Imagine, for example, if the subject's desire in Jiang et al.'s (2006) study was to purposefully control their saccadic eye movements to remain unchanged and neither attracted or repelled by the nude stimuli. Such an option was not available. They would be unable to control their reactions to *not* align with their self-reported gender preferences if they so chose. The subjects' reactions, therefore, seem to occur automatically, reflexively, and without any availability for guidance control by the person. Phillips (in Phillips and Block, 2016), builds on these points:

> Where attentional responses are completely stimulus-driven reflexes, operating entirely outside of voluntary control (e.g. Schoeberl et al. 2015), and possibly mediated by subcortical pathways (e.g. Mulckhuyse and Theeuwes 2010), I am unpersuaded that we must think of them as exercises of individual level agency. If they are not, we lack positive reason for thinking of the perceptual representations implicated by Jiang et al.'s data as constituting individual level perception (p. 183).

Finally, before moving onto section 3.2.3, consider indicator (v).

(v) Self-prompting, including the ability to act spontaneously and on cue:

It should be somewhat evident by this juncture how the alleged instances do not satisfy indicator (v). Nonetheless, it is worth reiterating that blindsight subjects often have to be persuaded into undertaking a 2afc task. Recall from Chapter One, for example, that the first extensively studied blindsight subject, DB, only completed the 2afc task after Weiskrantz et al. (1974) insisted that he take a guess whether the lines were horizontal or vertical; DB could not spontaneously tell them which direction the lines were. Moreover, it therefore seems that the content from the representations in blindsight subjects' blind field is not available for spontaneous self-prompting in such a way that they could use their blindsight in everyday life (namely, in a range of activities outside of a 2afc task). If they needed to make a phone call, for instance, and there was a mobile phone situated in their blind field, they would not spontaneously pick it up and make a call (without having pre-existing knowledge the phone was there). They would need to be prompted.

Similarly with DF. Recall again the example of putting mail into the slot. As we saw in indicators (i) and (iii), she was unable to prompt herself to do so spontaneously. Clark (2007) spells out:

> DF is famously unable to self-prompt the appropriate use of her own form-accommodating visuomotor skills. For example, she would not spontaneously attempt to post the letter through the slot, even though she had previously been told that a large prize would be given to the first successful posting … her experience does not present her with a world in which visual form information can provide the basis for deliberate ... action (p. 586).

Additionally, subjects under the effects of CFS are unable to use the content of their representations for self-prompted use. They cannot gather, for instance, the information from the nude stimuli or fearful faces and decide when, exactly, they will react to them. Instead, when and how they react is entirely at the mercy of subpersonal processes.

In sum, behaviour, as Kim (2011) iterates, is "the primary almost exclusive, evidence for the attribution of mental states to other beings with minds. Our knowledge of other minds depends primarily on observation of behaviour" (p. 87). If you saw me playing a game of soccer where I can use the content of my representations to pass the ball in a variety of directions depending on the options available to me—perhaps I decide to shoot when in a good position or pass to a teammate in a better one—you have good reason to think that I

am perceiving. You do so, because the content from my representations is evidently available for intentional coordination. However, if on the other hand you had to explain to me that there was a soccer ball directly in front of my eyesight and thus I was not able to spontaneously prompt myself to kick the ball, then my behaviour indicates that it has not adequately integrated into my cognitive economy in such a way that allows for intentionally coordinated guidance. The cases of blindsight, DF, and CFS subjects are just the same: they do not show *enough* of the above indicators for us to conclude that the content of their representations is available for intentional coordination.

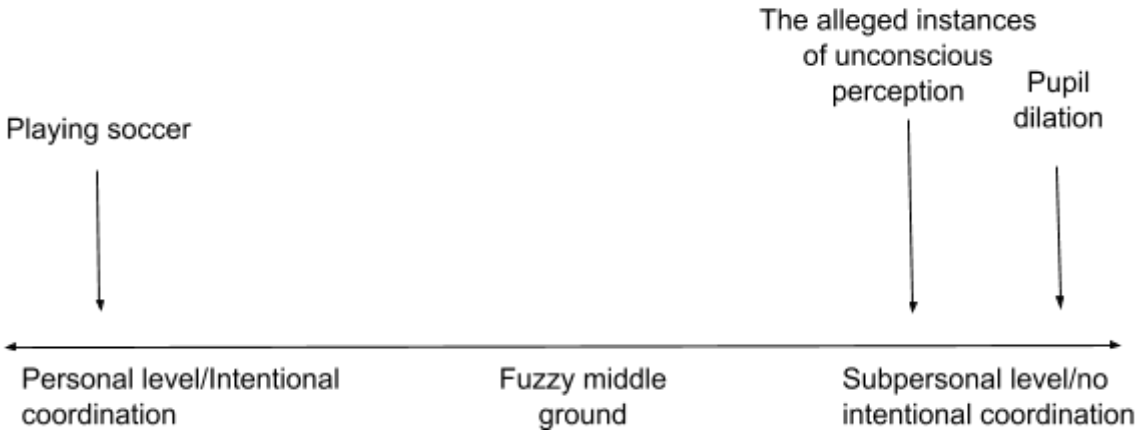
Nevertheless, the preceding analysis is not meant to be the final word and nor is it designed to be an exhaustive knockdown argument; I acknowledge that more might need to be spelt out. My aim, though, was to merely build enough of the defensive framework for Premise Three by showing that the alleged instances of unconscious perception very plausibly do not satisfy the demands for intentional coordination, and are thus not instances of perception.

3.2.3 Conclusion: The alleged instances of unconscious perception are not instances of perception.

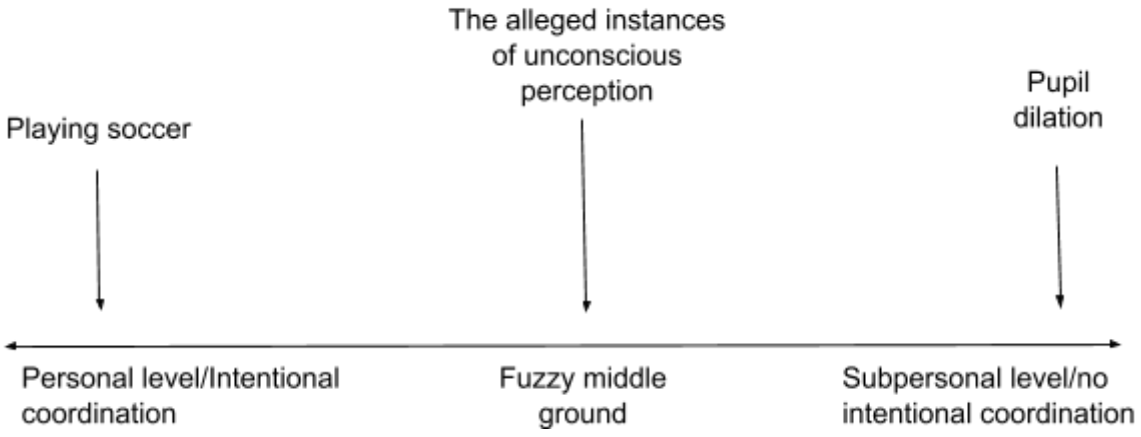
My claim is this: the alleged instances of unconscious perception are *indistinguishable* from subpersonal information processing. Indeed, in this chapter, I have presented what I take to be the most plausible criterion for perceptual attribution to the personal level. If we accept Premise Two—that intentional coordination is the necessary and sufficient criterion for personal level attribution—then I have examined in Premise Three how the alleged instances of unconscious perception are not personal level phenomena. If the alleged instances exercise any of the abilities of intentional coordination (even in a diminished manner), they do not exercise them to the extent demanded for personal level attribution. Therefore, the case for unconscious perception is undermined. Rather than being clear prima facie cases of unconscious perception, the alleged instances are, quite plausibly, clear prima facie instances of subpersonal sensory responses.

However, perhaps we should think of attribution on a spectrum. On one end, we have personal level phenomena where the contents of a representation are clearly available for intentional coordination (playing a game of soccer, for instance). On the other end, we have subpersonal phenomena which are clearly *not* available for intentional coordination (pupil dilation, for example). Plus, in the middle, we have a sort of fuzziness or a grey area where it

is not clear if the content is available for intentional coordination *or* if it is a subpersonal process (see Taylor, 2019, for a more in-depth discussion on this line of reasoning). I have argued for the substantial claim that the alleged instances of unconscious perception are not available for intentional coordination and are thus situated on the subpersonal end of the spectrum. However, perhaps one does not want to commit to the substantial claim. In that case, I hope to have demonstrated enough to show that one must, *at the very least*, be committed to the weaker claim which holds that it is unclear whether the alleged instances are personal or subpersonal level phenomena. They fall in the fuzzy penumbra between the two. So, according to the substantial claim I have defended, the personal/subpersonal level's spectrum of attribution looks like this:



Whereas according to the weaker claim, the spectrum of attribution looks like this:



The upshot, however, is that even if one admits to the weaker claim, then one has also admitted to plausible doubt whether the alleged instances of unconscious perception are attributable to the personal level. And with the personal level being a criterion of Burge's natural kind perception, we do indeed have reason to doubt there is unconscious perception of the same kind as conscious perception. Our analysis of whether there can be a mental state that is unconscious *and* fulfil Burge's criteria for perception, finishes looking like this:[31]

	Blindsight	**DF**	**CFS**
Is unconscious	**?**	**?**	✓
Is sensory	✓	✓	✓
Involves constancy	✓	✓	✓
Is individual-level	✗	✗	✗

With not all criteria 'ticked' on our report card, more work needs to be done to demonstrate we have unconscious perception. But what should the next move be? Is there a way to pave a path forward that moves the unconscious perception claim *beyond* a reasonable doubt? What avenues can an unconscious perception advocate explore? There are most likely several. For instance, perhaps the advocate can take advantage of the lack of a clear distinction between the personal and subpersonal levels by exploring the 'fuzzy' area between them. Can this fuzziness be sharpened to reveal further instances of unconscious perception? Moreover, if the distinction between perception and non-perception is fuzzy, does it also mean that instances in this fuzzy area are not different in *kind*? Maybe this could be an avenue to explore. However, if there is no clear line demarcating perception from non-perception, then instead of being a natural *joint* in nature, the difference between the kind perception and non-perception is more of a *curve*. This curve, I suggest, opens up a door for thinking of perception as a different sort of kind; namely, a manifest kind.

As flagged in Chapter One, Phillips (2018) also suggests that we can make progress in the unconscious perception debate by thinking of perception as a manifest kind. He argues that it is reasonable to think of perception as a manifest kind, but that if we do, there is no room for unconscious perception. I agree with Phillips that perception is quite plausibly a manifest

[31] Perhaps if one is only committed to the weaker claim, then instead of a '✗' next to individual-level for each alleged instance, we should place a '?'. Either way, as discussed in the main text, reasonable doubt remains.

kind, however, I *disagree* with him that there is no room for unconscious perception. In Chapter Four, I propose that *if* perception is a manifest kind, then perception appears to the third-person. Thus, consciousness is not required for perception. Indeed, contrary to Phillips, thinking of perception as a manifest kind can pave a way forward to finding unconscious perception.

Chapter Four: Paving a Path Forward

As it stands, there is a *live* possibility that we do not have unconscious perception. Chapters Two and Three gave us plausible reason for doubt. The analysis, however, worked under the framework that perception is a natural psychological kind. It demonstrated that *if* perception is a natural kind, then it is, at best, 'fuzzy' whether we have unconscious perception of the same kind as conscious perception. However, in this chapter, I shift frameworks from natural kinds and show that it is plausible to think of perception as a manifest kind instead. By doing so, I do not aim to settle the unconscious perception debate, I simply aim to pave a path forward which shows the direction an unconscious perception advocate *could* take in attempting to settle the debate. It is up to others to decide if they wish to walk the path.

So, first, I show why thinking of perception as a manifest kind is a reasonable position to take. Second, I argue that if perception is a manifest kind, then the manifest requirement is intentional coordination. I borrow from common-sense functionalism and draw on insights from studies on spiders and artificial intelligence (AI) to support this argument. Third, I show that if the key requirement for an instance of perception is a third-person manifestation rather than first-person phenomenal consciousness, then there is elbow room for an instance of unconscious perception. Consciousness is not required for perception; therefore, the unconscious perception claim has room to move. Indeed, for no apparent reason that I can discern, perception as a manifest kind is a neglected position and has thus received far too little attention in the literature; this chapter aims to rectify that.

4.1 Perception as a Manifest Kind

4.1.1 No Natural Carving

Firstly, independently of any argument *for* manifest kinds, I want to briefly put pressure on the claim that perception is a natural kind. As flagged at the end of Chapter Three, insofar as perception *is* a natural kind, then the kind seems to come in degrees or gradients, with the distinction between the kind perception and non-perception being fuzzy. However, if this is

the case, it poses a problem because a critical criterion for natural kinds is that they should be categorically distinct. As Bird and Tobin (2018) point out:

> There cannot be a smooth transition from one kind to another (Ellis 2001). For then the borderline between them could not be one drawn by nature, but one that is somehow or other drawn by us. In which case the kinds would not be genuinely natural (para. 14).

Natural kinds are meant to be non-arbitrary, objective distinctions carved up by joints in nature, not, that is, carved by us (Haslam, 2002). So if the demarcation is fuzzy, then any place within this fuzzy area where a line between kinds is drawn becomes artefactual.

Consider, for example, a problem adapted from Buckner (2016) involving Kind A and B, where the property *aa* identifies Kind A and the property *bb* identifies Kind B. We place Kind A and B on a spectrum; anything with the property *aa* is placed on one side and anything with *bb* on the other. For the most part, we find that there are clear members of each kind on either side of the spectrum. However, we also find that in the middle of the spectrum there are fuzzy instances where it is hard to distinguish if they should be categorised into Kind A or B. We might even find these fuzzy instances have the property *ab*. Nevertheless, it is our job to distinguish between kinds, so we decide to draw a line *somewhere* in this fuzzy area to *mostly* capture a distinction. However, the line is somewhat arbitrary, and it is definitely artefactual; it could have been further to the left or further to the right.

Moreover, regardless of whether we are in a perfect epistemic position, the cut will remain artefactual because there are not two distinct natural kinds to begin with; they blur into one another, so there was never a *natural* carving to be made. My claim, in other words, is that the problems faced by Kinds A and B are the same faced by the fuzzy distinction between perception and non-perception. In both cases, we have more than an epistemic problem, but a metaphysical one where there are not two natural kinds at play. The carving between perception and non-perception is fuzzy; thus, any carving eventually made by us is not natural, violating a fundamental criterion for natural kindhood.

Of course, philosophers can and do push back here. Boyd (1991, 1999), for example, offers a defence for the vagueness in natural kinds via Homeostatic Property Clusters (HPC), claiming that by adopting HPC, vagueness is a *feature* of natural kinds, rather than a

problem. Taylor (2019), moreover, does the same but with particular reference to perception. However, without pressing too far into the debate, I simply aim to point out that if one holds that natural kinds should indeed be naturally divided, then we have good reason for thinking perception is *not* a natural kind. If we take this problem seriously, then it opens the door into studying perception as a different type of kind, such as, I propose, manifest kinds. Unlike natural kinds, manifest kinds do not need to be carved by a *natural* joint in nature. Instead, a manifest kind is individuated by how it appears to us.

4.1.2 Manifest kinds

There are two essential details to understanding manifest kinds:

1. They have a manifest property; that is, they appear a certain way.
2. This appearance, or manifest property, is crucial to individuating the kind.

By 'manifest' or 'manifest property' I simply mean *to appear a certain way* with no underlying or hidden identity conditions.[32] A statue, for example, is a manifest kind because it appears a certain way to us and it appearing this way is crucial to individuating it from, say, a mere block of stone. Moreover, a gold ring, as Johnston (1997, p. 575) points out, is no longer a ring when rolled into a ball because appearing as a ring was crucial to individuating it from other kinds. Johnston (1997) further explains:

> The scientific image of the world is an image of how manifest kinds and their manifest instances are constituted so as to have the efficient causal powers that they do in fact have. Manifest kinds are thus satisfactorily located in a causal-explanatory view of the world. But, in general, they are not to be "reduced" by way of identification with chemical kinds.[33] *Their manifest form is at least as important to being what they are as is the matter that makes them up* (p. 580) [my emphasis].

So because the underlying causal properties do not identify manifest kinds, the block of stone which the statue is made from can never be used to *identify* a statue but can only *constitute* one when it meets the manifest requirement of appearing as a statue. Johnston

[32] Later, in section 4.2.3, I expand on what it means *to appear to us*.
[33] For our purposes substitute "chemical kinds" with "natural psychological kinds" and the point still stands.

(1997) gives a further example (which we first encountered in Chapter One) of H_20 and its manifest instances. Water vapour, for example, is a manifest kind whose appearance is a vapour (such as fog and mist), whereas snow is a kind whose appearance is powder. These manifest properties, rather than the underlying kind H_20, individuate the fog and snow. Manifesting into a fog is essential to being fog and manifesting into powder is essential to being snow.

However, a single quantity of H_20 can be powdery before it freezes into ice, then, the *same* quantity might liquify into water before finally turning into a vaporous form, so is this not a story of the same kind just manifesting into subkinds? No. If manifest kinds are to be the same, then their manifest instances must be the same (Johnston, 1997). While it is true that the underlying chemical kind H_20 remains the same quantity throughout, it is not true that its manifest instances remain the same. Ice cannot appear as fog and yet still be ice. A quantity of powder cannot appear as liquid water and yet still be snow. Indeed, if we use the underlying kind H_20 to identify snow *and* to identify water vapour, then, via the transitivity of identity, we seem to be committed to saying that snow must identify water vapour. This commitment is an untenable consequence. To avoid this absurdity, then, we should abandon the identity claim from the underlying natural kind H_20 and say that it cannot *identify* water vapour or snow but only *constitute* them if their manifest requirements are met. These manifest properties are the crucial identifiers of the kind.

To give another Johnstonian (1997) example, consider diamond, soot, and carbon. Carbon is the underlying kind of both diamond and soot, but again it would be untenable to say that carbon is used to identify *both* diamond and soot. Instead, we distinguish between diamond and soot because of their manifest properties. So the correct thing to say is that carbon only constitutes a diamond, and it only constitutes soot, if it meets the specific manifest requirements of each *because these manifest properties are crucial to individuating the kind.*

Perhaps, too, we can say something similar for perception. Let us say, for example, the underlying psychological kind for perception does involve Burge's (2010) criteria that we have been working with:

- is sensory
- involves constancy
- is individual-level.

However, just like a stone for a statue, H_20 for snow, and carbon for a diamond, perhaps Burge's underlying properties cannot be used to identify perception but can only constitute it if they manifest (appear to us) a certain way. Thus, even if we disregard perception's lack of natural carving and grant that it meets the natural kindhood criteria, perception still might be better understood as a manifest kind because it wears its identity conditions on its sleeve. In that case, we need to ask:

1. Does perception appear a specific way to us?

Yes. We saw in Chapter Three that perception appears a certain way via intentional coordination. Nevertheless, it is one thing to have a manifest property, and it is another thing for that property to be the key to individuating a kind. Most things have an appearance, but that does not make it a manifest kind. So the pertinent question is:

2. Is the appearance of intentional coordination crucial in individuating the kind perception?

Below in 4.2, I argue that the answer is yes. We have an instance of perception if (*and only if*) the manifest requirement of intentional coordination is met. Note that the central claim from Chapter Three and my claim here in Chapter Four are exclusive. Chapter Three was concerned with whether the alleged instances of unconscious perception fulfil Burge's criteria for the natural kind perception, and in doing so, showed that intentional coordination is the criterion for personal level attribution. However, Chapter Four is making the substantially different claim that the appearance of intentional coordination is the key to individuating the *manifest kind* perception. Accordingly, it is coherent for one to agree with me in Chapter Four without also needing to agree with Chapter Three. So long as the underlying conditions manifest into intentional coordination under the manifest kind framework—then whatever one thinks about individual-level attribution under Burge's framework—we have an instance of perception.

On top of this, note that when I say intentional coordination is the crucial identifier of perception, I am not just making an epistemic claim. Instead, I am claiming that we are genuinely obtaining access to perception's essential nature and job description. Intentional

coordination is a crucial property of the perceptual process; so we truly discover something about perception when we identify it. Consider Bennett and Hacker's (2003) observation:

> Possession of a sense-faculty is *manifest* in behaviour … the forms of behaviour that *manifest* possession of a given perceptual faculty consist in relative efficiency in discrimination, recognition, discernment, pursuit of goals and exploration of the environment (p. 127) [my emphasis].

I do not take Bennett and Hacker only to be claiming that behaviour enables us to *know* if someone is perceiving. Instead, there is a much more substantial metaphysical claim on the table that behaviour, or action, *is* the manifest expression of perception; that the function of perception manifests in the output of intentional coordination. This metaphysical claim is also my objective, and I distinguish it from any milder epistemic claims that might be a by-product. In other words, similar to when we see the colour green, we are not just gaining epistemic access to the fact that green looks different to yellow, instead we are also obtaining access to something crucial about *what green is*. So just as we know something intrinsic and essential to colour when we see it (Johnston, 1992), we also learn something intrinsic and essential to perception when we see it.

Indeed, green things, as Soames (2004) points out, are distinguished from yellow things, not necessarily because of any underlying causal structure, but because of how it appears. It would not *be* green as opposed to yellow if it did not appear as green. So to apply this to perception, the term 'perception' designates the property of the appearance that is causally responsible for the fact that instances of perception appear the same way to us (and different from non-perception instances). Hence, the predicate 'is perception' will apply to all and only those instances that have the property which causally explains why instances of perception look the same to us (and look different from instances of non-perception). In a nutshell, and to keep borrowing from Soames (2004), the way the property appears *determines* if an instance is a perceptual kind or non-perceptual kind. The manifest property, moreover, which the kind perception shares, and so *determines* its membership in the kind, and the property which non-perception instances lack, and so *determines* that it is not a perceptual kind, is, I submit, intentional coordination. In 4.2, I show why this is the case.

To sum up 4.1, though, there are two plausible reasons why perception should be considered a manifest kind (or, at least, not a natural psychological kind). One, it is not

categorically distinct from non-perception; the carving is not naturally made. Two, even if perception qualifies as a natural kind, we should drop the identity claim and say that it only constitutes perception if it meets the right manifest requirements. Either way, I have tried to show that we can reasonably understand perception as a manifest kind. Moving forward, therefore, I will be working under this framework and argue that intentional coordination is the crucial manifest property which individuates perception as a kind. However, to reiterate, Phillips (2018) claims that perception appears to the first-person qua consciousness. He says that perception is a manifest kind but that "such [manifest] requirements involve having a phenomenal nature. Hence all perception is conscious" (p. 478). So a large part of my task in what follows is to show why perception's essential identifier is a third-person appearance rather than a first-person one. The conclusion is a conditional one: *if* perception is a manifest kind, *then* the manifest requirement is intentional coordination which *therefore* provides elbow room for unconscious perception.

4.2 Manifest Kinds, Perception and Intentional Coordination

The purpose of 4.2 is to show that intentional coordination is the crucial identifier of the manifest kind perception. First, I will draw on insights from functionalism to help elucidate intentional coordination as perception's key role. If intentional coordination is perception's definitive functional output, then it is a good candidate for the property which individuates perception. Second, I will zoom in on common-sense functionalism and use examples from an optimally perceiving person, spiders, and AI to show that intentional coordination is what we already both commonly and implicitly take to be perception's identifier. Third, I will address some potential concerns about my claim by clarifying what it means to *appear to us*.

4.2.1 Intentional Coordination as the Common-Sense Function

Functionalism, as Armstrong (1968, 1981) spells out, describes a mental state as an internal state which is caused by sensory inputs and causes behavioural outputs. He says: "The concept of a mental state is primarily the concept *of a state of the person apt for bringing about a certain sort of behaviour* … [and secondarily] ... *apt for being brought about by a certain sort of stimulus*" (1968, p. 82) [Armstrong's emphasis]. So, the essential input and output of a mental state <input, output>, is what is apt for defining that state. For example:

<tissue damage, aversive behaviour> is apt for defining pain

<skin irritations, scratching> is apt for defining itch.

And I think we can say:

<external stimuli, intentional coordination> is apt for defining perception.

So for a person S to be in the mental state perception (that is, for S to instantiate perception) is just for S to be in some mental state (or to instantiate some property) with causal connections to the appropriate input (an external stimulus) and outputs (intentional coordination).[34] Moreover, there is a mesh between the input and output. They are tangled together as part of two sides of the *same* coin. Indeed, perception construed this way is similar to Millikan (1995) and Clark's (1997) conception when they state that rather than being a passive source of data, perception faces both ways at once: it says something about how the world is *and* prescribes possible actions.[35]

I am taking a role functionalist approach which should be distinguished from realiser functionalism. Role functionalism holds that two organisms, say a Martian and a human, are in the same mental state if their functional role is identical, regardless of what realises this function.[36] Kim (2011) further explains:

> Our focus should be on the roles that define these properties, and we should never forget that psychological-cognitive properties are role properties. Role functionalism identifies each mental property with being in a state that plays a specified causal role and keeps them clearly distinct from the physical mechanisms that fill the role, that is, the mechanisms that enable systems with the mental property to do what they are supposed to do (p. 187).

Role functionalism, then, is helpful for our purpose because we are in the business of trying to capture what is salient and definitive about perception. As Ross and Spurrett (2004)

[34] This is adapted to perception from Kim's (2011, p. 187) explanation of functionalism.
[35] Millikan and Clark mean slightly different things to Armstrong, but both points help to accurately capture what I am trying to say about perception.
[36] I do not have the time to engage properly with realiser functionalism. But, just briefly, it holds that we identify a mental property via the physical mechanism which realises it. So if perception in a Martian is realised by the function of, say, their antennas vibrating, and perception in a human is realised by the function of, say, our senses integrating stimuli, then the mental states are not the same. In this chapter I simply recognise that role functionalism is not universally accepted and pin my mast to it regardless. For a more detailed look at problems in role functionalism see Kim (1992).

further explain: “the very point of functionalism is to capture what is salient about what systems actually do, and how they interact, without having to get bogged down in micro-scale physical details” (p. 607), thus helping us to elucidate perception’s pivotal role and manifest property.

Furthermore, role functionalism is silent on how a function needs to be realised. For example, a heart is simply whatever organ pumps blood. Our human heart is anatomically different from a reptile or bird’s heart but because a heart’s job description (or definitive role) is to pump blood, then so long as an organ occupies this role, no matter how the role is realised, it remains a heart (Kim 2011, p. 131). Likewise, I propose, for perception. So long as the function occupies perception’s job description (or definitive role), then we have an instance of perception. Indeed, in the course of evolutionary history, to cope with hazards presented by their environment, organisms needed to develop a mental function which took the input from external stimuli and converted it into the output of intentionally coordinating the content of the stimulus. Briscoe (2014) makes a similar point in specific regards to vision:

> From an [sic] biological or evolutionary standpoint, it is reasonable to think that vision is *for* action, that its preeminent biological function is to adapt an animal’s bodily movements to the properties of the environment that it inhabits (p. 202) [Briscoe’s emphasis].

Patricia Churchland (1987), too, explains:

> looked at from an evolutionary point of view, the principle function of nervous systems is to get the body parts where they should be in order that the organism may survive (p. 548).

So, just like a heart, different species interacting with various environmental conditions, evolving independently of each other, have most likely developed distinctive realisers to achieve the definitive functional output. Some of these functional realisers might be conscious; others may not. Nevertheless, so long as the definitive function is realised, then regardless of how it is realised (whether consciously or unconsciously), we have an instance of perception. Indeed, I think a picture is beginning to form already that paints intentional coordination as the crucial identifier which individuates perception. But more needs to be said. Therefore, I will now zoom in and specifically look at common-sense functionalism. In

doing so, we will see that intentional coordination is what we already implicitly and commonly take perception's definitive function to be.

In zooming in on common-sense functionalism, however, I recognise that, just like role functionalism, it is not universally accepted. Again, though, I do not have the scope to address its objections in any worthwhile detail.[37] Instead, I offer two qualifiers to my use. One is to remember that my ultimate aim in 4.2 is to show that perception is a manifest kind identified by the property intentional coordination. I use functionalism—and in particular common-sense functionalism—to support this claim. Hence, one might hold that it is feasible I can achieve my aim independently of common-sense functionalism. If so, then we have an agreement on what is most important and a disagreement in the details. Yet in saying that, I think it necessarily follows from common-sense functionalism that perception is identified by intentional coordination. The second qualifier, therefore, is that if my ultimate claim is inseparable from common-sense functionalism, then my conclusion is conditional. *If* you agree that intentional coordination is perception's definitive function as identified by common-sense, *then* intentional coordination individuates perception. So, with that said, what exactly is common-sense functionalism?

In regards to perception, common-sense functionalism holds that:

> Perception is the inner mental state that plays the X role in the network of connections delivered by *common knowledge about the mind*. This network identifies a role for each mental state, and so perception is simply the state which occupies the role definitive of it.[38]

In other words, the mental state perception is just whatever function common-sense says is its definitive role. Moreover, I think we already implicitly and commonly take intentional coordination to be perception's definitive role. For example, perception can occupy many different functions. It can play the role, say, of being the topic of philosophical inquiry. It can also play the role of being the mental state which enables the organism to intentionally coordinate the content of their representations (and thus navigate their environment). Only one of these functional roles is definitive of the mental state perception, and I think we all

[37] For criticism on common-sense psychology see Paul Churchland (1981), and for a general discussion on the differences between common-sense functionalism and psycho-functionalism see Kim (2011, pp. 172 - 177) and Levin (2018).
[38] This conception of common-sense functionalism is adapted from Braddon-Mitchell and Jackson (2007, p. 53). Emphasis is mine.

commonly recognise this. If perception ceases to be a topic of philosophical inquiry, it would still be perception. However, if perception ceases to underpin an organism's ability to use the content of their representations to navigate their environment, then we have lost the definitive role that perception plays.[39] Nevertheless, even if we all commonly recognise this, why should we place so much importance on whatever common-sense folk psychology says is the definitive function?

Firstly, we accurately rely on our common-sense or implicit folk knowledge all the time in a variety of different ways. The classic example is grammar. One might understand a few basic rules about grammar, but not many of us would be able to write down a full account of the rules. Regardless, when reading, we commonly distinguish whether the writing is grammatically correct or not, even though we cannot always explicitly explain why. Instead, we rely on our implicit and common-sense understanding of grammar to evaluate the writing. Common-sense functionalism says that we do the same thing with the mind.[40]

We have an implicit, innate, and accurate ability to make predictions and explanations about the key features of people's mental states. When someone reaches for a drink, for example, most of us can identify that the reason he did so was that his definitive mental state was a desire to quench his thirst. Driving on a highway, moreover, would not be possible without an innate ability to understand other people's mental states. As Dennett (2009) extrapolates:

> We risk our lives without a moment's hesitation when we go out on the highway, confident that the oncoming cars are controlled by people who want to go on living and know how to stay alive under most circumstances. Suddenly thrust into a novel human scenario, we can usually make sense of it effortlessly, indeed involuntarily, thanks to our innate ability… (pp. 341-2).

Additionally, the chances, say, of predicting that people from all over the world would meet at a particular place at a particular time for a conference would be nothing short of a miracle without an inherent ability to understand mental states.[41]

---

[39] For a more in-depth, and early discussion, on how we decide what counts as most central to our common-sense understanding of mental states in general, see Lewis (1972).

[40] See Braddon-Mitchell and Jackson (2007, p. 62) for a longer discussion on the similarities of our implicit grasp of grammar and mental states.

[41] We might be able to write down some of what we are doing when we make these predictions, but the full account escapes us. See Braddon-Mitchell & Jackson (2007, pp. 48 - 64) for more discussion on our common ability to explain and make predictions. The example of predicting attendance at a conference comes from p. 63.

The examples could go on *ad nauseam*. However, the point is that we successfully rely on our implicit folk psychology all the time, to the extent that we *must* have some inbuilt ability to cotton onto what is important about mental states. Common-sense functionalism, therefore, says that because our folk psychology is so good, the definitive role in each mental state is simply whatever folk psychology says it is. Accordingly, the job of the grammar expert is to elucidate what we already implicitly understand about grammar. The psychologist, moreover, might have many different roles, but one of them is to help elucidate what we already implicitly understand about the mind. In a similar type of way, my job is to elucidate and make explicit what we already commonly and implicitly take the definitive role of perception to be. Our folk psychology has cottoned onto important functions in mental states, so if common-sense functionalism says that something plays the definitive role, we had better take note of what that is. Indeed, I think folk psychology already implicitly understands intentional coordination as the key feature of perception. I will now elucidate why this is so.

4.2.2 Elucidating our common-sense

As previously discussed, I am somewhat hinging my claim against Phillips' competing claim that says we use consciousness as the crucial identifier for perception. Accordingly, I want to distinguish between two questions:

1. Is the subject disposed to be able to intentionally coordinate their actions according to the content of their representations?

Call this the *action question*. How we answer the *action question* informs us whether the subject has met the manifest requirement *for perception*. Then, if we have answered 'yes' to the *action question*, we can ask question two:

2. Is the subject conscious?

Call this the *consciousness question*. How we answer the *consciousness question* informs us whether the instance of perception is conscious or unconscious. By distinguishing between these two questions, it will help clarify that consciousness is not our key manifest requirement for perception.

Therefore, let us answer the *action question* by evaluating the cases of an optimally perceiving person (Person A), spiders and AI. I will use the indicators of intentional coordination canvassed in Chapter Three to do so. Recall, the degree to which each indicator is satisfied and the quantity of indicators satisfied serves to enhance or weaken the case for perception. Also, an advantage of these indicators *for perception* is that they do not run into the same problems if they were *for consciousness*. Third-person measures of consciousness face criticism because they do not capture the subjectivity of consciousness. However, while consciousness necessarily entails subjectivity, I am merely putting forward indicators for perception. This distinction means that in answering the *action question*, we take no stance on consciousness and do not have the burden of capturing a subjective element; we can leave that for the *consciousness question*.

With that said, imagine Person A who can navigate his way through a unique environmental setting with many obstacles, moving targets, dangers, and random, unpredictable events. Let us say he is running through a busy city where cars, buses, pedestrians, streets, buildings, and noise all contribute to the complex environment. Can Person A intentionally coordinate? By answering yes, we will see that we already commonly and implicitly take the manifest property intentional coordination to be perception's crucial identifier and function.

> (i) Integration: Person A can integrate the content of his representations into the rest of his cognitive economy, enhancing his ability for intentional coordination. For example, when Person A sees white stripes across the road, he walks comfortably over them safe in the knowledge that the stripes signify a pedestrian's right of way. Indeed, even though he sees oncoming traffic on either side of the stripes, he also notices the cars are slowing down and integrates this with his knowledge that slowing down indicates braking. So, he is integrating the stimuli available to him with his working memory. Satisfying indicator (i) also helps him satisfy the others.
>
> (ii) Flexible behaviour: Person A can adapt to the ever-changing random and unique environment in a variety of different ways depending on the specific situation. For instance, he can stop suddenly at a traffic light, or slow down gradually if needed. Indeed, because of the contents of his representations, he is disposed to react in a variety of different ways depending on what any given situation demands, allowing him to make plans and problem-solve to achieve his goal flexibly.

(iii) Goal-directed behaviour: Person A has a goal of running to the other side of a busy block; he has implemented plans on how to achieve his goal by using the contents of his representation and knows that his actions are apt for obtaining the goal.

(iv) No contradiction or bypassing; the power of veto: Use of the mental content is not occurring independently of Person A. For example, when he sees a traffic light flash red or green, he can decide to stop or proceed his running when and how he would like. Perhaps when he first registers the stimulus of a road, he instinctively would like to run across it to achieve his goal, but because the light is red, he vetoes this initial reaction.

(v) Self-prompting: No one has told Person A that he must initiate his goal of reaching the other side of the block; nor has anyone told him that the light turned green and *now* is the time to move forward. He spontaneously and on cue reacts to new stimuli flexibly.

Person A is perceiving. We are justified in making this claim because perception's definitive functional role is apparent. Indeed, we have enough evidence that an instance of perception has manifested without needing to make any further claims about consciousness. If for whatever reason, when answering the *consciousness question*, we find that Person A is not conscious, (perhaps he is a philosophical zombie) we would not reverse our conclusion that he is perceiving. Person A perceives *independently* of any facts about whether or not he is conscious.

Moreover, concern about whether or not it is *actually* possible for Person A to be intentionally coordinating to the level he is without being conscious miss the point. I acknowledge that he almost certainly is conscious. However, the critical point here is that we have an instance of perception *regardless* of how we answer the *consciousness question*; perception hinges on the manifest property intentional coordination, not phenomenal consciousness. If it so happens that Person A is conscious, then that is fine, but it does not matter for now. Whether or not consciousness is involved is not crucial to whether it is perception; what is crucial is whether there is intentional coordination. *If* we take away consciousness, but the crucial role still manifests, that is, we still have intentional

coordination, then we would also still have an instance of perception. To sharpen this point, I will now analyse examples from spiders and AI.

4.2.2.1 Can Spiders Perceive?

Even with tiny brains, species of jumping spiders from the salticid subfamily Spartaeinae meet the manifest requirements for perception. By analysing spartaeines' manifestation of intentional coordination, it will become evident that we have a real-life scenario where we can answer 'yes' to the *action question* while remaining neutral on how to answer the *consciousness question*.

Cross and Jackson (2016) studied whether 15 spartaeines species spiders could use representations to guide their behaviour.[42] In the study, they positioned the spiders on a platform that looked out towards two distant towers. From this platform, the spiders could see that one of the towers had food, and the other did not. The spiders then had to navigate through a series of obstacles and detours towards the tower with food. However, once the spiders left the platform, they could no longer see the towers and none of the spiders tested had ever navigated the track before. The detour, moreover, took the spider on a route which forced them to turn 180 degrees away from the towers. The spiders were also encouraged to take routes to the wrong tower; their track took them past turn offs that led to the tower without food. However, Cross and Jackson (2016) found that regardless of the detours and obstacles, all 15 species of spartaeines spiders could consistently and successfully navigate their way to the tower with food. Despite their inability to see the tower on their route, their success in reaching the correct tower suggests they were not merely acting from a trial and error approach. That is, they were not bumping into walls until the right tower came back into sight, but, instead, they used *representations* to guide their behaviour. Indeed, their behaviour does not seem to be explained by a stimulus-driven response; after leaving the platform, they arrived at the correct tower by implementing a plan according to the contents of their visual representations.

[42] I have chosen to canvass this specific study, but the results from Cross and Jackson's (2016) study are not isolated, similar studies into the spartaeines species of *Portia* have yielded the same conclusion: these particular jumping spiders, which prey on other spiders, use representations to intentionally coordinate their behaviour. To read more into these similar studies see, for example, Tarsitano & Jackson (1997), Cross, Jackson & Pollard (2008), Cross & Jackson (2006, 2014, 2017). Below I also draw attention to a (1998) study by Stimson Wilcox and Jackson studying Portia in the field.

On top of this, similar studies by Stimson Wilcox and Jackson (1998) outside the laboratory confirm this conclusion. Stimson Wilcox and Jackson explain:

> During typical sequences in the field, Portia walks up the tree trunk toward A. appensa, stops, looks around, then goes off in a different direction, only later to come out above the web. There are usually vines and other vegetation near the tree and often some of the vegetation extends out above the web. After looking at the web and the surrounding environment (Jackson, 1992c; Jackson and Wilcox, 1993), Portia moves away, often going to where the web is completely out of view, crosses the vegetation and comes out above the web. From above the web, Portia drops on its own silk line alongside, but without touching, the web of the A. appensa. Then, when parallel with the spider in the web, Portia swings in to make a kill … appears that Portia takes deliberate detours to reach an optimal approach. Some of these detours observed in nature were over 1 metre in length and took over 1 hour to execute ... It is difficult to escape the conclusion that Portia solves detour problems … in its head, makes plans and then acts on these plans (pp. 422-3)

In other words, the variety of detour-related problems and tasks confronting Portia in the field would be even greater than in the lab and most likely includes problems that are even more complex. So let us apply the observations of spartaeines spiders onto our indicators of intentional coordination:

> (i) The representation of the tower with food stays integrated with their goal of reaching the food. For instance, Cross and Jackson were not able to trick the spiders into going down a decoy route because they had integrated their goal with their perceptual representation and knew that going down the wrong path will not help them reach their goal.
>
> (ii) The spiders do not seem to be exhibiting inflexibly instinct-driven behaviour. Instead, they use their representations to drive flexible behaviour that problem solves and reacts in a variety of context-sensitive ways depending on the content of their representations.

(iii) In the laboratory and the field, spartaeines spiders are directing their behaviour according to a goal. They are not just automatically reacting to new stimuli but using the content of their representations to pursue the goal of reaching their food/prey.

(iv) There does not seem to be any contradiction or bypassing when the spiders see the tower. They are not at the mercy of a sense driven stimulus but are controlling when and which route they use to reach their goal.

(v) Particularly in the field, Portia spiders seem to be self-prompting. They react according to cues by patiently waiting for minutes or hours until they prompt themselves according to the content of their representations that *now* is the time to act.

Spartaeines satisfy our indicators of intentional coordination, so we have an instance of perception. Moreover, as flagged, it is far from obvious that spiders are conscious; whether they are or not remains an open question. The upshot, therefore, is that in the case of spiders, we have an instance of perception which is hinged *not* on any facts or insights about whether they are conscious; instead, we have an instance of perception because the functional output of intentional coordination is satisfied. Indeed, we can remain neutral on whether they are conscious while still using a manifest property to individuate an instance of perception.

However, a would-be objector might still push back at this juncture. One could respond that because it remains an open question whether the spider is conscious, we are in a state of agnosticism, and agnosticism, the objector reasons, is not an argument *for* unconscious perception.[43] In other words, even if one accepts my manifest requirement and definitive function, I have still done nothing to show that we have unconscious perception. Indeed, the objector is right that I have not (yet) provided a *direct* argument for unconscious perception. However, the objector misses the point. The upshot of my discussion on spiders can be used in favour of an *indirect* argument. I have tried to show that the critical point drawn from the studies is that we should hinge an instance of perception onto a manifest requirement other than consciousness, *viz*. intentional coordination. Therefore, conceptually, we can have an instance that meets the requirements for perception and is not conscious. In 4.3, I say more about this conceptual elbow room where we might be able to find such instances

[43] This is similar to Phillips' (2018) objection when discussing spiders under a natural kind framework.

in humans and animals. First, however, I turn our attention to AI, which further builds on these points.

4.2.2.2 Can Artificial Intelligence (AI) perceive?

One of the lessons AI teaches us is that we should not be chauvinists about what realises the role of perception. As long as we can identify the manifest property, or the definitive function, it does not matter what realises the role. In what follows, therefore, I propose that specific types of AI can perceive. We will see that they do so because they satisfy the intentional coordination requirement for an instance of perception. Moreover, in analysing the case study of AI, it is noteworthy that while we can answer 'yes' for the *action question*, we can confidently answer 'no' for the *consciousness question*. Let us start by considering self-driving cars.

Alphabet Inc.'s company Waymo have developed self-driving cars which, much like Person A, can make models (representations) of their surrounding environments to safely navigate complex terrain, solve tasks and flexibly coordinate action. Waymo cars are equipped with technology which plays the functional role of our senses. For example, as Guizzo (2011) explains, the cars have laser-beams to build a detailed picture of the environment; radars to detect other objects' speed, and how far away these objects are; cameras to detect visual information such as traffic signals, global positioning systems (GPS), inertial measurement units, and wheel encoders. These 'senses' integrate with each other and with working memory. The laser-range finder, for example, generates a detailed 3D map of the environment which is integrated with high-resolution 'live' maps of the world; this allows the car to spot differences in the way the world appears to them 'live' with the way the world appeared in their memory. Comparing the new data with the old allows the car to distinguish between moving pedestrians and stationary objects. This integration (which satisfies (i) from our list of intentional coordination indicators) also helps the car to satisfy the flexibility (ii) and goal-directed behaviour (iii) indicators.

For example, as we saw with Person A, a busy city throws up many unique and variable situations requiring a high degree of flexibility for successful navigation. If someone quickly dashes across the road at the last second, the Waymo car can yield to the pedestrian (Fisac et al., 2019). It does not appear, moreover, to be an automated reflexion; the car can control itself to slow down calmly, swerve towards a different direction, or rapidly slam on the

brakes. Additionally, when driving through a large intersection, the car, for example, yields to other vehicles based on road rules. However, if other cars do not reciprocate, say they are aggressive drivers, it adapts to the conditions and advances a bit to show the other drivers its intention (Guizzo, 2011). This improvisation in novel and unexpected situations, along with the ability to act on the fly, indicates a high degree of flexible behaviour (ii) to satisfy its goals (iii).

The integration (i) also means that the use of the content does not bypass (iv) the car. The content does not stay isolated but can be used to veto a movement. When pedestrians dash across the street, for instance, the car uses the content to move in such a way that it does not hit the person. In other words, the car is not at the mercy of an automated system that keeps driving to pursue its goal of reaching a destination; instead, it can veto movements and use the content to serve its new goal of not hitting the person. However, the self-prompting indicator (v) is a little harder to evaluate. In one sense, the car seems to be pre-programmed by humans to behave and move the way it does, but in another sense, due to how flexible and spontaneous it is regarding its intended goal, it seems that the cars can self-prompt its movements in a variety of ways according to the content of its representations. For example, if it has scanned the street ahead and determined that the most efficient route to its destination is blocked with too many other cars or obstacles, it can prompt itself to enact another plan which takes it down a different route (Fisac et al., 2019). So, even if humans have initially programmed the cars with specific software, the programming, I think, seems to have given them the power to self-prompt a variety of flexible goal-directed movements. The only prompting by the human was in its initial programming; after that, it is hands-off.[44]

So what is the upshot of self-driving cars satisfying the indicators for intentional coordination? I propose that if it was not controversial to claim earlier that Person A was perceiving, then it is also not controversial now to claim the same for Waymo's self-driving cars. Functionally, in answering the *action question*, there are no relevant differences between them; that is, around very similar environments, both are using the content of their representations to coordinate their actions intentionally. Indeed, the only relevant difference between Person A and the self-driving cars is when we answer the *consciousness question*. Person A is almost certainly conscious, whereas the self-driving cars are almost certainly *not* conscious. How we answer the *consciousness question*, therefore, is not the definitive

[44] For further information on Waymo's self-driving cars, see their website (http://www.waymo.com).

identifier needed when ascribing a manifest instance of perception. Just like the spider, the AI are perceiving because the definitive property that determines whether a subject is perceiving, is apparent.

Perception might play many different roles, but only one is definitive. Perhaps most of the time, perception does play a role in providing content to conscious experience. Nevertheless, if we took away this role, that is, if a subject was not conscious and perception instead played the role of providing content only to unconscious representations, then we have not necessarily lost anything crucial to the function of perception. We have not lost anything because if the output of the function is still an ability to intentionally coordinate the content of the (now unconscious) representations, then we still have an instance of perception. What all instances of perception have in common is that they all serve a specific causal role distinctive of the kind perception; that distinctive causal role is intentional coordination.

In sum, we started this chapter by learning that manifest kinds meet two key conditions:

1. They appear a certain way; that is, they have a manifest property.
2. This appearance, or manifest property, is crucial to individuating the kind.

Perception appears a certain way to us, *qua* intentional coordination, thus satisfying (1). I then used common-sense functionalism to help elucidate that we already implicitly take intentional coordination to be the definitive role of perception. Consequently, it is crucial to individuating the kind and satisfying (2). Regardless of whether we find out the spiders, for example, are conscious or not, so long as we still have intentional coordination, then we have an instance of perception.

Indeed, just like it would be odd to say that something has the appearance of a statue without actually being a statue, I think it makes little sense to suppose that something (or someone) might have the appearance of intentional coordination without really being an instance of perception. Once we identify intentional coordination, we do not need to look past this manifest property into the underlying psychological kind. That is, we do not need to consult a psychologist or philosopher of perception to confirm that the instance is individual-level, constancy forming or sensory. We already have access to the nature of perception via its manifest property. So when we use the term ‘perception’, I think it is

plausible to say that we do not mean to refer to that thing which is individual-level and involves constancy that presents the appearance of perception; instead, we are referring to the manifest instance of perception because that appearance *is* definitive of perception.

I now want to address some potential concerns and clarify what it means for a property *to appear to us*.

4.2.3 What it Means to *Appear to Us*

One might agree that it is fine and well, as far as it goes, for perception to be a manifest kind, but there are problems in hinging perception onto a third-person manifestation. If the third-person is blind, for example, then does that mean perception never *appeared to us*? Moreover, what if I am colour-blind? Does that mean the grass is not green? Indeed, what if our view, say, of the statue is obstructed? In other words, am I erroneously relying on the faulty capabilities of the third-person to individuate perception? I do not think I am. Why? Because what it means to *appear to us*, I submit, simply refers to what it would mean to appear to *a suitably placed intentional systems expert who has all the relevant information*.[45] So by claiming that perception manifests to the third-person, I am saying that if the manifest property intentional coordination has appeared to the intentional systems expert, then we have an instance of perception.

However, what if the intentional systems expert, with all the relevant information at their disposal, thinks it is vague or fuzzy whether someone can intentionally coordinate and thus perceive? Does this imply that intentional coordination does not capture a distinction and thus cannot individuate kinds? No. Firstly, and to reiterate, manifest kinds do not need to be categorically carved by nature the same way natural kinds do; fuzziness does not violate any criterion of the kind. Secondly, as already mentioned, there *is* a fuzzy border between perception and non-perception, so if the manifest property intentional coordination can capture this fuzziness, it is a virtue of the approach. In other words, while the manifestation of intentional coordination remains crucial in individuating perception from non-perception, a fuzzy instance of intentional coordination is *also* crucial in identifying a fuzzy instance of

[45] This is similar to Johnston's (1998) explanation in regards to colour, where he says: "When an object has some color then standard subjects are under standard conditions disposed to see it as having that color" (p. 40). However, I am instead using Dennett's intentional systems theory as it more explicitly engages with what I have in mind in regards to folk-psychology. For a comprehensive explanation of the theory see Dennett (1987), but for a shorter piece see Dennett (2009).

perception. The border between perception and non-perception might be fuzzy, but intentional coordination captures this fuzziness.

Additionally, just because we might not be able to determine in a binary way perception from non-perception, it does not mean that we fail to delimit the concept in *any* way. Intentional coordination is not so vague that we allow any sensory episode to count as an instance of perception. The concept has enough meaning to still be a key identifier of perception, allowing us to grasp the distinction between perception and non-perception, *and* allowing us to capture the fuzzy middle ground between the two.

Okay, says the would-be objector, but what about the problem we encountered in Chapter Three regarding locked-in patients who do not have the motor ability to coordinate intentionally? Do they provide a counter-example which shows that we can have an instance of perception, and yet the function of perception does not appear to us?

Firstly, a locked-in patient's internal mental state of perception is almost certainly the same as anyone without locked-in syndrome. The difference is not in the mental function itself but in the pathway from the brain to the limbs. So perhaps we need to apply a *ceteris paribus* clause to perception. That is, all other things being equal, perception is the function which causes the behavioural output of intentionally coordinating the content from our representations after receiving the input from an external stimulus. This way, we still capture perception's definitive function, it is just that in the case of the locked-in patient, not all other things are equal, various means of intentionally coordinating the content are off the table because the pathway to the limbs has been disinhibited. But the mental function itself is still, as Armstrong (1981) might put it, suitably or aptly placed to play the definitive role. The content is still available for intentional coordination.

Secondly, in regards to how this function would *appear to us*, there most likely remain ways we can still observe the manifest property in an attenuated sense. For example, a locked-in patient can still intentionally coordinate their eye movements (Gallo & Fontanarosa, 1989); indeed, there is a difference between purposeful eye blinks and pupil dilation. When I concentrate on reading, I am intentionally coordinating my eyes across the page; the content of the representation is available to me as I purposefully guide my eyes at the appropriate pace. This eye movement, though, is quite different from mere stimulus-driven saccadic

reflexes. (see Kirchner & Thorpe, 2006 and de'Sperati & Baud-Bovy, 2008 for more on this distinction).

Alternatively, perhaps the patient's intentional coordination can appear via specific neuronal patterns in an fMRI machine. Consider, for instance, Owen et al.'s (2006) study of a 23-year-old female in a vegetative state. In the study, they tested the subject under two trials. In one, she was instructed to imagine playing tennis (motor imagery), and the other, to imagine visiting rooms in her home (spatial imagery). The study found that her blood oxygen level-dependent signal from the brain areas involved in motor imagery and spatial navigation were indistinguishable from 34 healthy volunteers under the same trials. Owen et al. (2006) explain:

> During the periods that she was asked to imagine playing tennis, significant activity was observed in the supplementary motor area. In contrast, when she was asked to imagine walking through her home, significant activity was observed in the parahippocampal gyrus, the posterior parietal cortex, and the lateral premotor cortex. Her neural responses were indistinguishable from those observed in healthy volunteers (p. 1402).

Owen et al. (2006) then go on to conclude that the ability to understand audible commands and cooperate with the trial in responding via brain activity shows "*a clear act of intention*, which confirmed beyond any doubt that she was consciously aware of herself and her surroundings" (p. 1402) [my emphasis]. I think we can object to it being "beyond any doubt" that the subject is conscious. Evidence of cognition does not definitively entail consciousness (see Naccache, 2006, Levy, 2008, Pennartz et al., 2019, and Shea & Bayne, 2010 for in-depth discussions on this point). We might also think that it is not a clear act of intentional coordination; however, the *appearance* of the correct neuronal patterns does seem to indicate that it is. I think she is probably in the functional state of perception because she uses the input of the instructions to produce an output of intentionally coordinating her response to the auditory content (but see Drayson, 2014, for an alternative account). However, I do not wish to give this a full examination because the point, in fact, is simply that a full examination *can be made*. Intentional coordination is still something that can appear to us.

Owen et al.'s (2006) study is also interesting because it helps clarify that the manifest property intentional coordination has a broader scope than we may have originally thought. It clarifies, for example, that we can intentionally coordinate the content of our representations via mental acts, not just motor acts. Owen, in a later (2013) study, makes this point:

> an "act" ... need no longer be a physical act in the traditional sense (e.g., the blink of an eye or the squeezing of a hand) but, with the aid of modern neuroimaging methods, can now be an act that occurs entirely within the brain itself—a "brain act," ... the ability to understand instructions and to carry out different mental tasks in response to those instructions, and therefore is able to exhibit willed, voluntary behavior in the absence of any overt action (p. 127).[46]

Thus, the upshot is this: so long as *in principle* there remain ways for perception's functional output to appear, then perception is still a kind individuated by the appearance of this output. That is, we have an instance of perception if and only if intentional coordination can appear*, in principle,* to the intentional systems expert with all the relevant information and right means of detection.

I suspect a lot more needs to be said on this, but at this stage, it seems like a plausible step in the right direction. For now, though, I want to turn our attention to the elbow room created for the unconscious perception advocate. Is there a small space, in regards to humans and animals, where we can answer 'yes' to the *action question* while answering 'no' to the *consciousness question*?

4.3 Elbow Room for Unconscious Perception

Phillips and I both agree that perception is plausibly understood as a manifest kind. However, we disagree on what the individuating manifest property is. He claims that the key property which individuates the kind perception is phenomenal consciousness. Whereas I

[46] As Mele (1997) also states: ''Some actions are 'overt': they essentially involve agents moving their bodies. Others are not: there are mental actions—for example, solving a chess problem in one's head, or deliberating about whether to accept a job offer'' (p. 231). However, again see Drayson (2014) who questions whether we can judge mental acts to be intentional acts. Indeed, she argues that it is not clear, for instance, whether Owen et al.'s (2006) study indicates an intentional mental act or a 'mere happening'. If she is right, though, then perhaps the upshot is that many mental acts fall in the fuzzy area between appearing as perception and non-perception and thus our approach to individuating perception has identified a fuzzy instance.

claim that intentional coordination individuates perception and thus makes it a manifest kind. If Phillips *were* right that perception manifests to the first-person, then we would arguably need to revise our folk psychological understanding of perception. The burden of proof is on him, I propose, to show why we should radically overhaul our folk understanding.

So, if consciousness is not required for perception to be manifest, then it creates some elbow room for the unconscious perception advocate to move. I now want to explore this elbow room. First, I will explain how manifest properties can come in different variations and degrees. Then, I will suggest that at the lower end of the manifest kind perception, we might be able to find an instance of intentional coordination without consciousness.

4.3.1 Manifest Kinds can Appear to Us in Different Degrees.

Every manifest property comes in different degrees and varieties. Snow, for example, might manifest in an abundance of powder or it might just manifest into one flake. A statue, moreover, might be carved to perfection or it might be poorly carved, defectively shaped and rough around the edges. Pain is also a manifest property that comes in different varieties. Breaking your ankle is more painful than breaking a nail, but breaking a nail can still satisfy the requirements for pain.

Furthermore, if an appearance individuates 'green' and is noticeably distinct from a different colour, say, 'blue', then we have an instance of a 'green' property. However, within this property, there may be many different shades and varieties (dark green, light green, lime green, pale green etc.), but if the instances are similar enough among themselves to still count as the same property green (and not blue), then they can individuate green. My point, in other words, is that in any given manifest kind, the variety of ways the individuating property can manifest is more finely grained than the kind itself.[47]

So, think again about perception. The appearance of intentional coordination individuates perception from non-perception, but within this manifest property, there can be a variety of ways and degrees it will appear to us. We can say that Person A from our earlier example of an optimal perceiver satisfies the manifest requirements for perception at a *gold standard*, but not every instance of intentional coordination will reach this gold standard. For example,

[47] Soames (2004) makes a similar point when he says: "it is important that kinds not be individuated as finely as the properties that determine them" (p. 166).

imagine Person B, who might maximally fulfil either *some* of the requirements for intentional coordination, or *most* of them to varying degrees, and thus shows *enough* of the indicators of intentional coordination to individuate perception without needing to reach the gold standard. Indeed, Person B is probably representative of what we are most of the time in our daily lives.

Moreover, think of Person C who fulfils each indicator of intentional coordination, but to a lesser degree than A or B. Their behaviour seems *sort of* flexible but there are certain options and evaluations that they do not seem capable of. For example, they cannot self-prompt, but once prompted, their behaviour shows signs of flexibility. Intentional coordination might still appear, but only minimally. Next, imagine Person D. This person has to be asked to react to a stimulus; they are not able to self-prompt or spontaneously coordinate the content of his representations. Moreover, they might be somewhat goal-orientated but only under specific task conditions, and it seems habitual rather than a planned goal, meaning they do not have a wide array of behavioural responses to the content. I think Person D, in fact, might closely resemble someone like DF in that it is clear the person has *some* sensory response to stimuli but because of their diminished ability, they fall short of being an instance of intentional coordination.

To spell this line of reasoning out in a bit more detail, consider the below graded indicators of perception scale (see next page). Each number represents a score, and I suggest that perception has appeared to us *qua* intentional coordination if the sum of all the scores reaches *anywhere* between 16 and 26:

	0	2	4	6
Integration	No integration	Little indication of integration	Good reason to think the stimulus is integrated	Appropriately integrated with no reasonable doubt
Flexible behaviour	No flexible behaviour	Minimal indication of flexible behaviour	Mostly seems to display flexible behaviour	Reaches maximum capacity of flexible behaviour
Goal-directed behaviour	No goal-directed behaviour	Minimal indication of goal-directed behaviour	Mostly seems to display goal-directed behaviour	Full goal-directed behaviour
No contradiction or bypassing; the power of veto	Contradiction and bypassing	Mostly bypassed but small indicators of veto power	No contradiction or bypassing	N/A
Self-prompting	No self-prompting	Needs to be encouraged for self-prompting but once encouraged can coordinate	Is fully self-prompting	N/A

I recognise that a scale like this for perception may look very similar to a scale which one might use for consciousness.[48] However, consciousness is not my target here. It is important to be clear that the above scale is only interested in perception; it is solely concerned with the *action question* and thus designed to help us understand when perception has appeared to us and how it can appear in different varieties. For instance, Person A would most likely record a score of 26 out of 26 and therefore reach a gold standard of intentional coordination. However, as just canvassed, an instance of perception does not have to reach this gold standard to still appear to us. If someone, say, only scores 4's for 'integration',

[48] See, for instance, the Glasgow Coma Scale (GCS) which is a neurological test performed by physicians. The GCS is used as a test for consciousness on patients suffering from head injuries, trauma or are in intensive care. See Teasdale and Jennett (1974) for an early introduction of the scale, and Teasdale et al. (2014) for a more recent review.

'flexible behaviour' and 'goal-directed behaviour', and 2's for 'no contradiction or bypassing' and 'self-prompting', then they will record a 16 and still pass the minimum threshold for an instance of perception. Perception has not appeared to us at the gold standard but has nevertheless still appeared.[49] So the point is that any score between 16 and 26—while appearing in different degrees and varieties—is still an instance of perception appearing to us. With that said, consider my next point.

4.3.2 Perception Without Consciousness?

It is plausible that around the *minimum* threshold for appearing as perception (16 out of 26), we might find an instance of perception that is not conscious. Indeed, while it is reasonable to think that the gold standard of intentional coordination might turn out to also be a hallmark of consciousness, by the same token, it is also reasonable to think that the lower the level of intentional coordination, the lower the likelihood of it being accompanied by consciousness.[50] A score of 16 for perception could only, I think, *suggest* consciousness without being decisive or indisputable. Therefore, it opens up some elbow room for an instance to satisfy the minimum threshold for perception and yet not be conscious. Indeed, we saw in Chapter Two that behaviour, or action, is not an exclusive measure of consciousness; action may also measure unconscious cognitive abilities.

Further, we saw above with AI (and depending on one's theory of consciousness perhaps spiders too) that the function of perception does not rely, *as such*, on consciousness to realise the role of intentional coordination. It is conceptually possible to separate intentional coordination from consciousness, and if perception appears to the third-person, then consciousness is not required for an instance of perception. In other words, while a score of 16 (for perception) may often be indicative of consciousness, it is not a surefire marker of consciousness in humans and animals.

Note that I am not claiming that a score of 16 for perception *is* unconscious. I recognise that in most cases, it is reasonable to think that a score of 16 for perception is also conscious. However, I am putting pressure on the thought that it is *always* conscious. Are there any instances of humans or animals passing the minimum threshold for perception and yet doing

[49] I suggest a score of 16 because I think it mostly captures a distinction between intentional coordination appearing verses not appearing, but am open to the minimum threshold being revised.
[50] For more on the reasoning behind the gold standard perhaps being a hallmark for consciousness, see, for instance, Shea and Bayne (2010) who call it 'the argument from volition', and also see (for example) Dretske, 2006, Bayne, 2013, and Pennartz et al., 2019.

so in the absence of consciousness? I think this question remains somewhat open. Even though most instances of intentional coordination may be accompanied by consciousness, it is not clear that it *always* is; particularly down the lower end of the threshold. It is worth looking for exceptions. If there are any exceptions, then we have unconscious perception.

Finally, it is worth clarifying that I am *not* claiming to have finished the story of making explicit what we implicitly understand about intentional coordination and perception. But I am claiming to have taken a step in the right direction towards explicitly fleshing out what we mean by perception and intentional coordination. It was clear at the end of Chapter Three that more needed to be said to advance the unconscious perception claim; the burden of proof was on the unconscious perception advocate to move the argument beyond a reasonable doubt. This chapter has tried to pave a path forward for the advocate to show how this move might be made. I have shown that perception is plausibly understood as a manifest kind and that *if* it is a manifest kind, then contrary to Phillips (2018), instead of appearing to the first-person via consciousness, it appears to the third-person via intentional coordination. Within this third-person appearance we might find an instance of perception without consciousness; it is an open question whether we will or not. I finish, however, by applying Armstrong's (1981) caveat to this chapter: "I have done no more than sketch a programme … There are all sorts of expansions and elucidations to be made, and all sorts of doubts and difficulties to be stated and overcome. But I hope I have done enough to show ... an exciting and plausible intellectual option" (p. 302). This is also my hope.

Conclusion

I started this thesis by asking whether humans and animals have unconscious perception. The short answer given was 'perhaps'. Now, four chapters later, we are in a position to understand the long answer of *why* this is the case. There were four significant steps made along the way.

First, I presented three alleged instances of unconscious perception—blindsight, DF's visual form agnosia, and CFS—to stand as a representative sample for *any* alleged instance. I created an unconscious perception report card involving Burge's criteria for perception to evaluate whether these instances are indeed both unconscious and perception. If no alleged instance recorded a 'tick' next to all criteria on the report card, then there is reasonable doubt whether we have unconscious perception.

Second, I argued that it is an open question whether blindsight subjects and DF are unconscious of the stimuli they report not being aware of. Whenever a subject exhibits d' > 0 and yet does not report awareness of the stimulus, subjective measures sustain the belief they are not conscious of the stimulus. However, as we saw, subjective measures face the problem of the criterion and do not reliably distinguish between degraded conscious information and unconscious information processing. It is a *live* possibility, therefore, that blindsight subjects and DF are conscious of stimuli and yet fail to report this dim and degraded conscious information. The upshot, though, is not that they *are* conscious of a stimulus, it is that we are not in the position to know—beyond a reasonable doubt—either way. On the other hand, insofar as a CFS subject exhibits d' = 0, then I suggested that even though there is doubt that they are not conscious of the stimuli, perhaps it is beyond a reasonable doubt. Recording d' > 0 is not a reliably exclusive measure of consciousness, nor is it likely to be infallibly exhaustive, however, I think it is fair to say that it is *reliably* exhaustive. Accordingly, CFS recorded a 'tick' on our unconscious perception report card for the criterion 'is unconscious', whereas blindsight and DF's status remains an open question.

Nevertheless, in Chapter Three, I bracketed concerns about whether blindsight subjects and DF are unconscious of the stimuli and asked: if each alleged instance is unconscious, are they even instances of perception? By using Burge's criteria for perception as the framework for my analysis, I argued that the alleged instances do not satisfy a key criterion of his: they fail to be attributable to the individual-level. I argued that attribution requires the content of a

representation to be available for intentional coordination whereas the alleged instances do not meet this requirement. I defended the 'substantial claim' which holds that rather than being prima facie examples of perception, the alleged instances are prima facie examples of subpersonal information processing. However, I also tried to show that even if one is not willing to commit to my substantial claim, one must, at the very least, commit to the 'weaker claim' which holds that it is 'fuzzy' or unclear whether the alleged instances are personal or subpersonal level phenomena. Either way, even if one only concedes to the weaker claim, then one has admitted to reasonable doubt over whether the alleged instances are indeed instances of perception. Our unconscious perception report card finished looked like this:

	Blindsight	**DF**	**CFS**
Is unconscious	**?**	**?**	✓
Is sensory	✓	✓	✓
Involves constancy	✓	✓	✓
Is individual-level	✗	✗	✗

With no alleged instance recording four ticks, more needs to be said to move the unconscious perception claim *beyond* a reasonable doubt.

In Chapter Four, therefore, I paved *a* path forward for the unconscious perception debate by proposing that perception is plausibly understood as a manifest kind individuated by the appearance of intentional coordination *to the third-person.* I argued that *if* perception is a manifest kind, then contrary to Phillips (2018), there is room for unconscious perception. I used common-sense functionalism to support my claim that intentional coordination and *not* consciousness is the key manifest requirement for perception. The upshot is that because an instance of perception hinges on its appearance to the third-person, conceptually, there is elbow room for an instance of unconscious perception. It is within this elbow room that I suggest we look.

Of course, there no doubt needs to be much more said on perception as a manifest kind. For example, at the end of Chapter Four I fleshed out a 'graded indicator of perception scale' that I put forward as a helpful guide to determining when perception has appeared to us.

This guide, though, is not meant to be the final word on the topic; instead, it is an invitation for more discussion and refinement. Some might think, for instance, that my third-person appearance is on the right track but do not want the threshold for this appearance to be as high as I have set it. If so, that would be interesting to explore. Can the threshold for an instance of perception appearing to us be lower than what I have set it and yet still capture the difference between perception and non-perception? Are there weaker forms of intentional coordination that can be endorsed? I suggested that a score of 16 out of 26 on the scale is the minimum threshold for intentional coordination—anything below that and it starts to get fuzzy—however, can more be said about this fuzzy area between perception and non-perception? If we can sharpen the fuzziness, for instance, then perhaps we can lower the minimum threshold and discover instances of unconscious perception.

Other questions also remain. For example, we saw in Chapter Two that behaviour is not an exclusive measure of consciousness; so, how do we verify when—and to what extent—it tracks consciousness? Moreover, what if someone satisfies the minimum threshold for intentional coordination and yet does not report conscious awareness of the stimuli, at what point, if ever, does their report become reliable? How do we verify their report? Can subjective measures ever become reliably exhaustive? These are pertinent questions and honing the answers to them are vital to moving the unconscious perception claim beyond a reasonable doubt.

Indeed, this thesis has put forward a common-sense way of individuating perception; a way which can help us make significant strides forward in understanding when, or to what degree, any current or future alleged instances of unconscious perception is actually perceiving. It is reasonable to understand perception as a manifest kind, and I have shown that if it is a manifest kind, then it appears to the third-person and thus consciousness is not required for perception. Therefore, *the* principal question which this thesis leaves us with is: can we find an instance of intentional coordination—that is, perception *appearing* to us—without consciousness? If we can answer ‘yes’ beyond a reasonable doubt, then humans and animals do have unconscious perception.

References

Armstrong, D. M. (1968). *A materialist theory of the mind*. London, New York,: Routledge & K. Paul Humanities Press.

Armstrong, D. M. (1981). *The nature of mind, and other essays*. Ithaca, N.Y.: Cornell University Press.

Azzopardi, P., & Cowey, A. (1997). Is blindsight like normal, near-threshold vision? *Proc Natl Acad Sci U S A, 94*(25), 14190-14194. doi:10.1073/pnas.94.25.14190

Azzopardi, P., & Cowey, A. (1998). Blindsight and visual awareness. *Conscious Cogn, 7*(3), 292-311. doi:10.1006/ccog.1998.0358

Azzopardi, P., & Hock, H. S. (2011). Illusory motion perception in blindsight. *Proc Natl Acad Sci U S A, 108*(2), 876-881. doi:10.1073/pnas.1005974108

Bayne, T., & Spener, M. (2010). Introspective Humility. *Nous*, 1-22.

Bayne, T. (2013). Agency as a marker of consciousness. In A. Clark, J. Kiverstein, & T. Vierkant (Eds.), *Decomposing the Will* (pp. 160-180). United Kingdom: Oxford University Press.

Bengson, J. J., & Hutchison, K. A. (2007). Variability in response criteria affects estimates of conscious identification and unconscious semantic priming. *Conscious Cogn, 16*(4), 785-796. doi:10.1016/j.concog.2006.12.002

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell Pub.

Berger, J. (2014). Mental States, Conscious and Nonconscious. *Philosophy Compass, 9*(6), 392-401.

Berger, J., & Mylopoulos, M. (2019). On Scepticism about Unconscious Perception. *Journal of Consciousness Studies, 26*(11-12), 8-32.

Bird, A., & Tobin, E. (2018). Natural Kinds. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 ed.). \url{https://plato.stanford.edu/archives/spr2018/entries/natural-kinds/}: Metaphysics Research Lab, Stanford University.

Blackmore, S. J. (2004). *Consciousness: an introduction*. Oxford ; New York: Oxford University Press.

Blake, R., & O'Shea. (2009). Binocular Rivalry. In Encyclopedia of Neuroscience (pp. 179-187). Elsevier.

Block, N. (2005). Two neural correlates of consciousness. *Trends Cogn Sci, 9*(2), 46-52. doi:10.1016/j.tics.2004.12.006

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav Brain Sci, 30*(5-6), 481-499; discussion 499-548. doi:10.1017/S0140525X07002786

Block, N. (2008). Consciousness and cognitive access. *Proceedings of the Aristotelian Society, 108*(1pt3), 289-317.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends Cogn Sci, 15*(12), 567-575. doi:10.1016/j.tics.2011.11.001

Block, N. (2014). Seeing-As in the Light of Vision Science. *Philosophy and Phenomenological Research, 89*(1), 560-572.

Block, N. (2016). The Anna Karenina Principle and Skepticism about Unconscious Perception. *Philosophy and Phenomenological Research, 93*(2), 452-459.

Blythe, I. M., Kennard, C., & Ruddock, K. H. (1987). Residual vision in patients with retrogeniculate lesions of the visual pathways. *Brain, 110 (Pt 4)*, 887-905. doi:10.1093/brain/110.4.887

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *An International Journal for Philosophy in the Analytic Tradition, 61*(1), 127-148. doi:10.1007/BF00385837
Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141-185): MIT Press.
Braddon-Mitchell, D., & Jackson, F. (2007). *The philosophy of mind and cognition* (2nd ed.). Malden, MA: Blackwell Pub.
Brentano, F. (1874). *Psychology From an Empirical Standpoint*: Routledge.
Briscoe, R. (2014). Spatial Content and Motoric Significance. *Avant: Trends in Interdisciplinary Studies*(2), 199-216.
Brown, R. (2014). Consciousness doesn't overflow cognition. *Front Psychol, 5*, 1399. doi:10.3389/fpsyg.2014.01399
Buckner, C. (2016). Transitional Gradation in the Mind: Rethinking Psychological Kindhood. *British Journal for the Philosophy of Science, 67*(4), 1091-1115.
Burge, T. (2010). *Origins of Objectivity*: Oxford University Press.
Campion, J., Latto, R., & Smith, Y. M. (1983). Is blindsight an effect of scattered light, spared cortex, and near-threshold vision? *Behavioral and Brain Sciences, 6*(3), 423-486.
Carruthers, P. (1996). *Language, thought, and consciousness : an essay in philosophical psychology*. Cambridge England ; New York, NY, USA: Cambridge University Press.
Carruthers, P. (2017). Block's Overflow Argument. *Pacific Philosophical Quarterly*, 65-70.
Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Percept Psychophys, 36*(4), 387-395. doi:10.3758/bf03202793
Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Can J Psychol, 40*(4), 343-367. doi:10.1037/h0080103
Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy, 78*(2), 67-90. doi:10.2307/2025900
Churchland, P. S. (1987). Epistemology in the Age of Neuroscience. *The Journal of Philosophy, 84*(10), 544-553. doi:10.5840/jphil1987841026
Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge, Mass.: Cambridge, Mass. : MIT Press.
Clark, A. (2001). Visual Experience and Motor Action: Are the Bonds Too Tight? *Philosophical Review, 110*(4), 495.
Clark, A. (2007). What Reaching Teaches: Consciousness, Control, and the Inner Zombie. *The British Journal for the Philosophy of Science, 58*(3), 563-594. doi:10.1093/bjps/axm030
Clark, A. (2009). Perception, action, and experience: Unraveling the golden braid. *Neuropsychologia, 47*(6), 1460-1468. doi:10.1016/j.neuropsychologia.2008.10.020
Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends Cogn Sci, 15*(8), 358-364. doi:10.1016/j.tics.2011.06.008
Cowey, A. (2010). The blindsight saga. *Exp Brain Res, 200*(1), 3-24. doi:10.1007/s00221-009-1914-2
Cross, F. R., & Jackson, R. R. (2006). From eight-legged automatons to thinking spiders. In K. Fujita & S. Itakura (Eds.), *Diversity of Cognition: Evolution, Development, Domestication and Pathology* (pp. 188 - 215). Kyoto: Kyoto University Press.
Cross, F. R., Jackson, R. R., & Pollard, S. D. (2008). Complex display behaviour of Evarcha culicivora, an East African mosquito-eating jumping spider. *New Zealand Journal of Zoology, 35*(2), 151-187. doi:10.1080/03014220809510112
Cross, F. R., & Jackson, R. R. (2014). Specialised use of working memory by Portia africana, a spider-eating salticid. *Anim Cogn, 17*(2), 435-444. doi:10.1007/s10071-013-0675-2

Cross, F. R., & Jackson, R. R. (2016). The execution of planned detours by spider-eating predators. *J Exp Anal Behav, 105*(1), 194-210. doi:10.1002/jeab.189
Cross, F. R., & Jackson, R. R. (2017). Representation of different exact numbers of prey by a spider-eating predator. *Interface Focus, 7*(3), 20160035. doi:10.1098/rsfs.2016.0035
D'Aloisio-Montilla, N. (2019). Phillips on Unconscious Perception and Overflow. *Philosophia (United States), 47*(3), 649-662. doi:10.1007/s11406-018-0021-7
Danckert, J., & Rossetti, Y. (2005). Blindsight in action: what can the different sub-types of blindsight tell us about the control of visually guided actions? *Neurosci Biobehav Rev, 29*(7), 1035-1046. doi:10.1016/j.neubiorev.2005.02.001
Davies, M. (1995). Consciousness and the varieties of aboutness. In C. Macdonald & G. Macdonald (Eds.), *Philosophy of Psychology: Debates on Psychological Explanation* (pp. 356 - 392). Oxford: Blackwell Publishers.
de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *Neuroreport, 10*(18), 3759-3763. doi:10.1097/00001756-199912160-00007
de'Sperati, C., & Baud-Bovy, G. (2008). Blind saccades: an asynchrony between seeing and looking. *J Neurosci, 28*(17), 4317-4321. doi:10.1523/JNEUROSCI.0352-08.2008
Debner, J. A., & Jacoby, L. L. (1994). Unconscious perception: attention, awareness, and control. *J Exp Psychol Learn Mem Cogn, 20*(2), 304-317. doi:10.1037//0278-7393.20.2.304
Dehaene, S. (2014). *Consciousness and the brain: deciphering how the brain codes our thoughts*. New York, New York: Viking.
Dennett, D. C. (1969). *Content and consciousness*. New York,: Humanities Press.
Dennett, D. C. (1987). *The Intentional Stance*: MIT Press.
Dennett, D. C. (1991). *Consciousness explained*. Boston: Boston : Little, Brown and Co.
Dennett, D. C. (2001). Surprise, surprise. *Behavioral and Brain Sciences, 24*(5), 982-982.
Dennett, D. C. (2009). *Intentional Systems Theory*.M. In Beckermann, A. B. P., and Walter, S, (Ed.), *The Oxford Handbook of Philosophy of Mind* (pp. 339-50): Oxford University Press.
Dienes, Z., & Seth, A. K. (2010). Measuring any conscious content versus measuring the relevant conscious content: comment on Sandberg et al. *Conscious Cogn, 19*(4), 1079-1080; discussion 1081-1073. doi:10.1016/j.concog.2010.03.009
Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *J Exp Psychol Gen, 127*(3), 286-303. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9742717
Drayson, Z. (2014). Intentional Action and the Post-Coma Patient. *An International Review of Philosophy, 33*(1), 23-31. doi:10.1007/s11245-013-9185-8
Drayson, Z. (2017). Psychology, Personal and Subpersonal. In T. Crane (Ed.), *Routledge Encyclopedia of Philosophy*.
Dretske, F. (2006). Perception without awareness. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual Experience* (pp. 147--180): Oxford University Press.
Elton, M. (2000). Consciousness: Only at the personal level. *Philosophical Explorations, 3*(1), 25-42.
Elton, M. (2003). *Daniel Dennett: Reconciling Science and Our Self-Conception*: Polity.
Eriksen, C. W. (1960). Discrimination and learning without awareness: a methodological survey and evaluation. *Psychol Rev, 67*, 279-300. doi:10.1037/h0041622
Fisac, J. F., Bronstein, E., Stefansson, E., Sadigh, D., Sastry, S. S., & Dragan, A. D. (2019). Hierarchical game-theoretic planning for autonomous vehicles. In (Vol. 2019-, pp. 9590-9596).

Fischer, J. M., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*: Cambridge University Press.
Fisk, G. D., & Haase, S. J. (2006). Exclusion failure does not demonstrate unconscious perception II: Evidence from a forced-choice exclusion task. *Vision Research, 46*(25), 4244-4251. doi:10.1016/j.visres.2006.08.023
Frankfurt, H. G. (1978). The Problem of Action. *American Philosophical Quarterly, 15*(2), 157-162.
Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: an experimental framework. *Trends Cogn Sci, 3*(3), 105-114. doi:10.1016/s1364-6613(99)01281-4
Gallo, U., & Fontanarosa, P. (1989). Locked in Syndrome - Report of a case. *Am. J. Emerg. Med., 7*(6), 581-583.
Goldman, A. (1979). What is Justified Belief? In G. S. Pappas (Ed.), Justification and Knowledge (pp. 1 -25). Dordrecht: Reidel.
Goodale, M. A., & Milner, A. D. (2004). *Sight unseen : an exploration of conscious and unconscious vision*. Oxford ; New York: Oxford University Press.
Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
Guizzo, E. (2011). How Google's Self-driving Car Works. Retrieved from https://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works
Haslam, N. (2002). Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry, and Psychology, 9*(3), 203-217.
Hohwy, J. (2011). Phenomenal Variability and Introspective Reliability.(Report). *Mind & Language, 26*(3), 261. doi:10.1111/j.1468-0017.2011.01418.x
Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. Behavioral and Brain Sciences, 9(1), 1-23.
Hornsby, J. (2000). Personal and sub-personal; A defence of Dennett's early distinction. *Philosophical Explorations, 3*(1), 6-24.
Hurley, S. L. (1998). *Consciousness in action*. Cambridge, Mass. London: Cambridge, Mass. London : Harvard University Press.
Irvine, E. (2009). Signal detection theory, the exclusion failure paradigm and weak consciousness--evidence for the access/phenomenal distinction? *Conscious Cogn, 18*(2), 551-560. doi:10.1016/j.concog.2008.11.002
Irvine, E. (2012). Old Problems with New Measures in the Science of Consciousness. *The British Journal for the Philosophy of Science, 63*(3), 627-648. doi:10.1093/bjps/axs019
Irvine, E. a. (2013a). *Consciousness as a Scientific Concept A Philosophy of Science Perspective* (1st ed. 2013. ed.): Dordrecht : Springer Netherlands : Imprint: Springer.
Irvine, E. (2013b). Measures of Consciousness. *Philosophy Compass, 8*(3), 285-297. doi:10.1111/phc3.12016
Irvine, E. (2019). Developing Dark Pessimism Towards the Justificatory Role of Introspective Reports. *Erkenntnis*, 1-26. doi:10.1007/s10670-019-00156-9
Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proc Natl Acad Sci U S A, 103*(45), 17048-17052. doi:10.1073/pnas.0605678103
Jiang, Y., & He, S. (2006). Cortical Responses to Invisible Faces: Dissociating Subsystems for Facial-Information Processing. Current Biology, 16(20), 2023-2029.
Johnston, M. (1992). How to speak of the colors. *Philosophical Studies, 68*(3), 221-263.
Johnston, M. (1997). Manifest kinds. *Journal of philosophy, xciv*(11), 564-583.

Johnston, M. (1998). Are manifest qualities response-dependent? *Monist, 81*(1), 3-43.
Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition, 79*(1-2), 89-113. doi:10.1016/s0010-0277(00)00125-6
Kentridge, R. W., & Heywood, C. A. (1999). The status of blindsight: Near-threshold vision, islands of cortex & the Riddoch phenomenon. *Journal of Consciousness Studies, 6*(5), 3-11.
Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (1999). Effects of temporal cueing on residual visual discrimination in blindsight. *Neuropsychologia, 37*(4), 479-483. doi:10.1016/s0028-3932(98)00084-0
Kentridge, R. W. (2011). Attention without awareness: A brief review. In C. Mole, D. Smithies, & W. Wu (Eds.), *Attention: Philosophical and Psychological Essays* (pp. 228-246). Oxford: Oxford University Press.
Kentridge, R. W. (2015). What is it like to have type-2 blindsight? Drawing inferences from residual function in type-1 blindsight. *Conscious Cogn, 32*, 41-44. doi:10.1016/j.concog.2014.08.005
Kim, J. (1992). Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research, 52*(1), 1-26. doi:10.2307/2107741
Kim, J. (2011). *Philosophy of mind* (3rd ed. ed.). Boulder, CO: Boulder, CO : Westview Press.
Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res, 46*(11), 1762-1776. doi:10.1016/j.visres.2005.10.002
Klotz, W., & Neumann, O. (1999). Motor Activation Without Conscious Discrimination in Metacontrast Masking. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 976-992. doi:10.1037/0096-1523.25.4.976
Koch, C. (2004). *The quest for consciousness: a neurobiological approach*.Denver, Colo. : Roberts and Co.
Kouider, S., Dehaene, S., Jobert, A., & Le Bihan, D. (2007). Cerebral bases of subliminal and supraliminal priming during reading. *Cereb Cortex, 17*(9), 2019-2029. doi:10.1093/cercor/bhl110
Kriegel, U. (2013). A hesitant defense of introspection. *Philosophical Studies, 165*(3), 1165-1176.
Levin, J. (2018). Functionalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018 ed.). https://plato.stanford.edu/archives/fall2018/entries/functionalism/: Metaphysics Research Lab, Stanford University.
Levy, N. (2008). Going Beyond the Evidence. *The American Journal of Bioethics, 8*(9), 19-21. doi:10.1080/15265160802318261
Lewis, D. K. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy, 50*(3), 249-258.
Lin, Z. & Murray, S.O. (2014) Priming of awareness or how not to measure visual awareness. Journal of Vision, 14(1), 27.
Mele, A. R. (1997). Agency and mental action. *Nous-Supplement: Philosophical Perspectives, 11*(1), 231-249.
Merikle, P. M. (1984). Toward a definition of awareness. *Bulletin of the Psychonomic Society, 22*(5), 449-450. doi:10.3758/BF03333874
Merikle, P. M., & Joordens, S. (1997). Measuring unconscious influences. In *Scientific approaches to consciousness.* (pp. 109-123). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
Millikan, R. G. (1995). Pushmi-pullyu representations. *Philosophical Perspectives, 9*, 185-200.

Naccache, L. (2006). Is She Conscious? *Science, 313*(5792), 1395-1396. doi:10.1126/science.1132881
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83*(October), 435-450.
Overgaard, M. (2006). Introspection in science. *Conscious Cogn, 15*(4), 629-633. doi:10.1016/j.concog.2006.10.004
Overgaard, M., Fehl, K., Mouridsen, K., Bergholt, B., & Cleeremans, A. (2008). Seeing without Seeing? Degraded Conscious Vision in a Blindsight Patient. *PLoS One, 3*(8), e3028. doi:10.1371/journal.pone.0003028
Overgaard, M. (2015). The challenge of measuring consciousness. In M. Overgaard (Ed.), *Bevioural Methods in Consciousness Research* (pp. 7 - 20). Oxford: Oxford University Press.
Overgaard, M., & Sandberg, K. (2012). Kinds of access: different methods for report reveal different kinds of metacognitive access. *Philos Trans R Soc Lond B Biol Sci, 367*(1594), 1287-1296. doi:10.1098/rstb.2011.0425
Overgaard, M. (2017). The Status and Future of Consciousness Research. *Front Psychol, 8*, 1719. doi:10.3389/fpsyg.2017.01719
Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science, 313*(5792), 1402. doi:10.1126/science.1130197
Owen, A. M. (2013). Detecting consciousness: a unique role for neuroimaging. *Annu Rev Psychol, 64*, 109-133. doi:10.1146/annurev-psych-113011-143729
Pennartz, C. M. A., Farisco, M., & Evers, K. (2019). Indicators and Criteria of Consciousness in Animals and Intelligent Machines: An Inside-Out Approach. *Front Syst Neurosci, 13*, 25. doi:10.3389/fnsys.2019.00025
Perenin, M. T., & Rossetti, Y. (1996). Grasping without form discrimination in a hemianopic field. *Neuroreport, 7*(3), 793-797. doi:10.1097/00001756-199602290-00027
Peters, M. A. K., Kentridge, R. W., Phillips, I., & Block, N. (2017). Does unconscious perception really exist? Continuing the ASSC20 debate. *Neurosci Conscious, 2017*(1), nix015. doi:10.1093/nc/nix015
Phillips, I. (2016). Consciousness and Criterion: On Block's Case for Unconscious Seeing. *Philosophy and Phenomenological Research, 93*(2), 419-451. doi:10.1111/phpr.12224
Phillips, I., & Block, N. (2016). Debate on unconscious perception. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception* (pp. 165 - 192). New York: Routledge.
Phillips, I. (2018). Unconscious Perception Reconsidered.(Report). *Analytic Philosophy, 59*(4), 471. doi:10.1111/phib.12135
Pisella, L., Grea, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., Rossetti, Y. (2000). An 'automatic pilot' for the hand in human posterior parietal cortex: toward reinterpreting optic ataxia. *Nat Neurosci, 3*(7), 729-736. doi:10.1038/76694
Ramsøy, T., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences, 3*(1), 1-23. doi:10.1023/B:PHEN.0000041900.30172.e8
Reingold, E. M., & Merikle, P. M. (1990). On the Inter-relatedness of Theory and Measurement in the Study of Unconscious Processes. *Mind & Language, 5*(1), 9-28. doi:10.1111/j.1468-0017.1990.tb00150.x
Ross, D., & Spurrett, D. (2004). What to say to a skeptical metaphysician: A defense manual for cognitive and behavioral scientists. *Behavioral and Brain Sciences, 27*(5), 603-627.
Sanders, M. D., Warrington, E. K., Marshall, J., & Wieskrantz, L. (1974). "Blindsight": Vision in a field defect. *Lancet, 1*(7860), 707-708. doi:10.1016/s0140-6736(74)92907-9

Schroeter, F. (2004). Endorsement and Autonomous Agency. *Philosophy and Phenomenological Research, 69*(3), 633 - 659.
Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review, 117*(2).
Schwitzgebel, E. (2019). Introspection. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019 ed.). https://plato.stanford.edu/archives/win2019/entries/introspection/: Metaphysics Research Lab, Stanford University.
Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*: Cambridge University Press.
Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn Sci, 12*(8), 314-321. doi:10.1016/j.tics.2008.04.008
Shea, N., & Bayne, T. (2010). The Vegetative State and the Science of Consciousness. *Br J Philos Sci, 61*(3), 459-484. doi:10.1093/bjps/axp046
Shea, N. (2013). *Neural Mechanisms of Decision-Making and the Personal Level*: Oxford University Press.
Snodgrass, M. (2002). Disambiguating conscious and unconscious influences: do exclusion paradigms demonstrate unconscious perception? *Am J Psychol, 115*(4), 545-579. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12516528
Snodgrass, M., & Lepisto, S. A. (2007). Access for what? Reflective consciousness. *Behavioral and Brain Sciences, 30*(5-6), 525-526.
Soames, S. (2004). Knowledge of manifest natural kinds. *Facta Philosophica, 6*(2004), 159-181.
Soto, D., Mantyla, T., & Silvanto, J. (2011). Working memory without consciousness. *Curr Biol, 21*(22), R912-913. doi:10.1016/j.cub.2011.09.049
Spener, M. (2015). Calibrating Introspection. *Philosophical Issues, 25*(1), 300-321.
Spener, M. (forthcoming). Consciousness, introspection, and subjective measures. In U. Kriegel (Ed.), *The Oxford Handbook of the Philosophy of Consciousness*: Oxford University Press.
Sperling, G. (1960). The Information Available in Brief Visual Presentations. *74*(11), 1-29. doi:10.1037/h0093759
Stein, T. (2019). The breaking continuous flash suppression paradigm: Review, evaluation, and outlook. In G. Hesselmann (Ed.), Transitions between Consciousness and Unconsciousness. Routledge
Stimson Wilcox, R., & Jackson, R. R. (1998). Cognitive abilities of araneophagic jumping spiders. In I. Pepperberg, A. Kamil, & R. Balda (Eds.), *Animal Cognition in Nature* (pp. 411 - 433). San Diego: Academic Press.
Stoerig, P., & Cowey, A. (1997). Blindsight in man and monkey. *Brain, 120 (Pt 3)*, 535-559. doi:10.1093/brain/120.3.535
Tarsitano, M. S., & Jackson, R. (1997). Araneophagic jumping spiders discriminate between detour routes that do and do not lead to prey. *Anim. Behav., 53*, 257-266.
Taylor, H. (2019). Fuzziness in the Mind: Can Perception be Unconscious? *Philosophy and Phenomenological Research*, <xocs:firstpage xmlns:xocs=""/>. doi:10.1111/phpr.12592
Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet, 2*(7872), 81-84. doi:10.1016/s0140-6736(74)91639-0
Teasdale, G., Maas, A., Lecky, F., Manley, G., Stocchetti, N., & Murray, G. (2014). The Glasgow Coma Scale at 40 years: standing the test of time. *Lancet Neurol, 13*(8), 844-854. doi:10.1016/S1474-4422(14)70120-6

Timmermans, B., & Cleeremans, A. (2015). How can we measure awareness? An overview of current methods. In M. Overgaard (Ed.), *Behavioural Methods in Consciousness Research* (pp. 21-48). Oxford: Oxford University Press.
Trevethan, C. T., Sahraie, A., & Weiskrantz, L. (2007). Can blindsight be superior to 'sighted-sight'? *Cognition, 103*(3), 491-501. doi:10.1016/j.cognition.2006.04.011
Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nat Neurosci, 8*(8), 1096-1101. doi:10.1038/nn1500
Visser, T. A., & Merikle, P. M. (1999). Conscious and unconscious processes: the effects of motivation. *Conscious Cogn, 8*(1), 94-113. doi:10.1006/ccog.1998.0378
Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain, 97*(4), 709-728. doi:10.1093/brain/97.1.709
Weiskrantz, L. (1986). *Blindsight : a case study and implications*. Oxford [Oxfordshire] : New York: Oxford Oxfordshire : Clarendon Press New York : Oxford University Press.
Weiskrantz, L. (1997). *Consciousness lost and found : a neuropsychological exploration*. Oxford ; New York: Oxford University Press.
Wu, W. (2018). The Neuroscience of Consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018 ed.). https://plato.stanford.edu/archives/win2018/entries/consciousness-neuroscience/: Metaphysics Research Lab, Stanford University.
Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion, 7*(4), 882-886. doi:10.1037/1528-3542.7.4.882
Yang, E., Brascamp, J., Kang, M. S., & Blake, R. (2014). On the use of continuous flash suppression for the study of visual processing outside of awareness. *Front Psychol, 5*, 724. doi:10.3389/fpsyg.2014.00724