An argument against the conjunction of direct realism and the
standard causal picture

P. H. Griffiths

A thesis submitted for the degree of Doctor of Philosophy

Department of Philosophy and Art History

University of Essex

April 2016

Abstract

Recent work in defence of direct realism has concentrated on the representationalist and disjunctivist responses to the arguments from illusion and hallucination, whilst relatively little attention has been given to the argument from causation which has been dismissed lightly as irrelevant or confused. However such charges arise from an ambiguity in the thesis which is being defended and the failure to distinguish between metaphysical and epistemological issues and between factual and conceptual claims. The argument from causation, as an argument against the conjunction of metaphysical direct realism and an explanation of the perceptual process in terms of a naturalistically understood causal chain of events, has not been answered in the philosophical literature. Moreover when the process of perception is fleshed out in terms of contemporary cognitive science, the difficulties are compounded. Neither representation-friendly mainstream cognitive science, nor representation-averse radical embodied cognitive science, is compatible with a theory of perception which is at the same time both direct and robustly realist.

Contents

Acknowledgements

The University of Essex Philosophy Department has provided a very stimulating and supportive environment in which to do this research. The organizational framework provided by the supervisory board meetings and the feedback and suggestions have been invaluable; in this respect particular thanks are due to David McNeil and Dan Watts. Thanks also to all those who have contributed to the discussion of my work in the Graduate Seminars/Conferences and Departmental Research Colloquia.

Most of all, I must thank Wayne Martin for his support and supervision throughout my period of study. Without exception, our supervisory meetings and ad hoc discussions have been productive and intellectually stimulating. Moreover, not only has Wayne guided me through the various stages of this project, he has developed my knowledge and understanding of a wider range of issues.

Introduction

At the beginning of the twentieth century the majority of philosophers dismissed the direct realist theory of perception as hopelessly naïve, but by the final decades it had become the orthodoxy. However, this rehabilitation of direct realism has concentrated on the representationalist (aka intentionalist) and disjunctivist responses to the arguments from illusion and hallucination, whilst the argument from causation, understood as an argument "based on the physical and physiological mechanism of perception" (Armstrong, 1961, p. 140) has received relatively little attention and has been dismissed lightly. Moreover, although the notion of direct perception is rejected by mainstream cognitive science, the claim that direct realism is nevertheless compatible with a naturalistic explanation of the perceptual process is implicit within much of contemporary analytic philosophy. Significantly then, our central thesis is that:

Metaphysical direct realism[1] is incompatible with an explanation of the perceptual process in terms of a naturalistically understood causal chain of events.

If, in accordance with this thesis, there is a radical incongruity between our pre-critical intuition of the immediacy of perception and an explanation of the perceptual process in terms of a causal chain of events, we might expect to experience a feeling of genuine puzzlement when we try to reconcile the two. It is this which Valberg describes in *The Puzzle of Experience* (Valberg, 1992).

Towards the end of his detailed exploration of the tension between the seeming immediacy of perception and a causal explanation of the perceptual process, Valberg discusses the "everyday

---

[1] We define "metaphysical direct realism" as the conjunction of a metaphysical directness-claim, that the immediate object of awareness in the veridical case is one-and-the-same-as the physical object, and a realism-claim, that the physical world has an existence that is not in any way dependent on it being cognized. See Section 1.2 for detailed discussion.

puzzlement" which this line of thought invokes. He considers the situation in which, innocent of the philosophical debate, we might review the facts of visual perception:

> …light is reflected by the external thing, it impinges on the eye, and so on. How if we are being innocent, would we complete the story? How, in particular would we conceive of the upshot of the causal process, the part that we vaguely refer to as 'experience'? (Valberg, 1992, p. 204)

After ruling out philosophically sophisticated responses, he suggests that:

> We take the experiential upshot of the causal process to be the presence of an external object. Things happen in the brain and then, as a result, the world is present. This is our innocent view of the matter, but there is something puzzling about it. Our own innocent conception of the experiential outcome of the causal process is incomprehensible to us. (p. 207)

It seems that there is no way that what happens in the brain could cause an external thing to be present. "But it is stronger than that: there is *no way that there could be a way*. There is, as we might say, no 'conceivable' way" (p. 213) Moreover, as Valberg emphasizes, the puzzlement arises from the seeming impossibility of reconciling our pre-critical intuitions with the causal picture. The puzzlement has nothing to do with external world skepticism or the possibility of illusion or hallucination. If we were not wedded to the causal picture, we would not be puzzled:

> Puzzlement arises **only** [emphasis added] when we combine the fact of the world's presence with the causal picture of experience; when, that is, we attempt to view the fact of presence as caused by what is happening in the brain. (p. 215)

Here is the puzzle which motivates this dissertation and lies at the core of the argument from causation. Significantly, the argument from causation first came to the fore in the seventeenth century with "the atomist picture of the physical world [in which] perception depends on causal chains which link the perceiver and the object they perceive." (Robinson, 1994, p. 84) As Armstrong puts it:

> The argument bases itself on the scientific discovery that, before we can perceive anything, a chain of processes must begin in the object, travel through our sense-organs, and reach the brain. (Armstrong, 1961, p. 141)

Without such a causal picture the puzzlement does not arise and the argument from causation cannot be framed.   We dub this the "standard causal picture"; a more precise definition will follow.

We can introduce the three positions in the philosophy and science of perception in terms of their respective responses to this puzzle. Mainstream cognitive science and indirect realist philosophy accept the standard causal picture, as fleshed out by twenty first century science, but maintain that it (the picture) is incompatible with the notion of direct perception; the puzzle is resolved at the expense of our pre-critical intuitions of the directness of perception. Gibsonian cognitive science supports the notion of direct perception, but only by questioning the metaphysical assumptions which underpin the standard causal picture; the puzzle is resolved at the expense of the standard causal picture. The direct realism of contemporary analytic philosophy stands apart, by retaining both the standard causal picture and direct perception. On this account, there is no puzzle to be resolved; the standard causal picture is in fact compatible with our pre-critical intuitions.

Correspondingly, there are three positions with respect to the argument from causation. Mainstream cognitive science and indirect realism accept that the argument is valid and sound. Gibsonian cognitive science maintains that the argument is valid but, since it assumes the standard causal picture, unsound. Direct realism accepts the standard causal picture, but rejects the argument as invalid; it is claimed that the argument is confused or irrelevant.

Although it will sometimes be useful to distinguish between, say, the philosophical and scientific literature, the continuity of the science and philosophy of perception is an underlying theme of this dissertation. Indeed no such distinction would have been recognized by the Early Moderns and it would be impossible to categorize much of the recent work, say that in "neurophilosophy" (Churchland, 2002) or "radical embodied cognitive science" (Chemero, 2009). Moreover we will

take issue with those who, implicitly or explicitly, distance the philosophy of perception from the debate within cognitive science. In this respect we sympathize with Hirst who, writing at a time when Linguistic Philosophy was at its height, argues that philosophers should continue to aim, in the tradition of Descartes, Locke, Berkeley etc, to "press into philosophical service the discoveries of science in order to achieve the aim of a full understanding of the nature of man and his relation to the world." (1959, p. 135)

In this dissertation we bring together the philosophy and cognitive science to examine all three responses to the argument from causation. Indeed, here, on the ground of Valberg's puzzle and the argument from causation, the philosophy and science are inextricably linked. Only philosophically sophisticated cognitive science, or scientifically savvy philosophy, call it what you may, has the resources to address the problem.

Drawing on these resources we defend our central thesis by arguing that the argument from causation has been dismissed too lightly in the philosophical literature, and does in fact present serious and unresolved difficulties for metaphysical direct realism; much of the literature is compromised by a failure to distinguish between metaphysical and epistemological claims and between factual and conceptual claims. Moreover these difficulties are compounded when the argument is fleshed out in terms of contemporary cognitive science. Neither mainstream Helmholtzian representation-friendly cognitive science nor the representation-averse Gibsonian alternative is compatible with a theory of perception which is both direct and robustly realist.

Understood in these terms, the issues raised by the argument from causation are at the core of our understanding of the world and our relationship to it. We will be at pains to reject deflationary construals of the argument and the issues arising. As Noë and Thompson put it in the introduction to a collection of papers on the philosophy and science of perception:

> It would be difficult to overstate the degree to which the problems dealt with by the writers in this book are alive and unresolved. …. what is at stake is ultimately one's understanding of consciousness itself and one's conception of one's place in the natural world. (2002, p. 12)

Overview

This is not a dissertation that begins with a literature review. Rather the defence of our central thesis will unfold in continuing dialogue with recent developments in both the philosophy of perception and cognitive science. Chapters 1 and 2 are primarily concerned with the argument from causation in the context of the debate between direct and indirect realism in the philosophical literature. Chapter 3 introduces the contra-parallel debate within cognitive science, and Chapter 4 focuses on the implications of mainstream representation-friendly theories of perception. (Appendix A addresses the avowedly radical representation-averse theories.)

Chapter 1 has three main objectives. Firstly, to position the argument from causation with respect to all four traditional arguments against direct realism[2] and the late twentieth century defences thereof[3]. Secondly, to disambiguate the target of the argument by distinguishing between directness v realism, metaphysical v epistemological, factual v conceptual and negative v positive claims; much of the literature defending direct realism is seriously compromised by failure to address these crucial distinctions. Lastly, to refine the argument from causation as an argument against the conjunction of the standard causal picture and metaphysical direct realism factually construed.

We do not claim that metaphysical direct realism is false, only that it is incompatible with the standard causal picture. To this end, we formalize the logic of the argument in terms of five premises and a general argument schema, particular instantiations of which draw on the standard causal picture to claim particular facts concerning the immediate object of awareness which,

---

[2] The arguments from illusion, hallucination, causation and science.

[3] Primarily the disjunctivist and representationalist versions of direct realism.

contrary to metaphysical direct realism's identity claim, do not hold of the normal physical object.[4] We maintain that, in this form, the argument is valid and the premises are robust though not entirely uncontroversial.

Chapter 2 addresses the objections raised in the philosophical literature starting with Armstrong's critique in his seminal defence of direct realism (1961). We find three types of objection to the traditional argument from causation: type-1, that the argument is self-refuting since the conclusions undermine the premises; type-2, that the considerations, with respect to the physical and physiological mechanisms of perception, adduced by the argument are irrelevant to the standing of direct realism; and type-3, that the argument depends on the controversial Phenomenal Principle.

Although type-1 objections present serious difficulties for the traditional version of the argument which takes the standard causal picture as a premise, they do not apply to a version which claims only that the standard causal picture and the metaphysical directness-claim cannot both be true. Similarly, although type-2 objections might serve in the defence of the epistemological directness-claim or some *strictly-conceptual* [5]claim, they have no purchase against an argument which is clearly targeted on the metaphysical directness-claim factually construed. Type-3 objections are rarely raised, since most direct realists would be unwilling to follow Armstrong in rejecting the notion of an immediate object of awareness. However such an objection could be maintained by the influential minority who espouse a thoroughgoing reductive representationalist (aka intentionalist) analysis of veridical perception.

 At this point we have defused the objections to the argument from causation on the basis of which it has been dismissed lightly as irrelevant or confused. Indeed, far from being the poor

---

[4] Here we draw on Smith's (2002, p. 8) analysis of the general structure of the arguments against direct realism.

[5] That is a claim which places a principled restriction on the type of data which is allowed count for or against. See Section 2.2.3.

relation in the family of four traditional arguments, it presents the most serious challenge to direct realism. However the thoroughgoing representationalists' type-3 objection remains to be addressed. Moreover an overarching type-2 irrelevance-objection might be pressed on the grounds that "direct realism" is a *strictly-conceptual* thesis which stands independently of **any** considerations with respect to the mechanism of perception. Both objections come at the cost of seriously weakening the thesis which is being defended and would be contentious within contemporary analytic philosophy. To break the impasse we need to extend the scope of our discussion.

Chapter 3 introduces the contra-parallel debate within cognitive science. Here Gibson's avowedly revolutionary approach (1968) (1979) challenges the Helmholtzian orthodoxy (Gregory, 2002) (Rock, 2004) of indirect perception. Although Gibson's theory of direct perception remains highly controversial, recent developments with respect to embodiment, computation and representation have been taken to lend support to his ideas, notably by proponents of "radical embodied cognitive science" (Chemero, 2009). We do not take sides in the debate. The key point is that, when theories of direct perception are maintained, it is at the expense of rejecting much of orthodox cognitive science and espousing a radical alternative which is at odds with the standard causal picture.

The question now arises as to the connection between these developments and the debate between direct and indirect realism within analytic philosophy. We maintain that they are inextricably linked. That perception involves inference is a key tenet of Helmholtzian cognitive science, whereas Gibsonian cognitive science aligns with the epistemological directness-claim. Moreover by virtue of rejecting the standard causal picture, and thus disarming the argument from causation, Gibsonian cognitive science lends support to the metaphysical directness-claim, albeit at the expense of a potential conflict with the realism-claim. The significant issues which

stand between direct and indirect theories of perception can only be progressed by an alliance between philosophy and cognitive science; indeed, in some quarters the alliance is well established and any demarcation would be hard to find. (Wheeler, 2005), (Chemero, 2009), (Churchland, 2012)

In Chapter 4 we draw on mainstream representation-friendly cognitive science to develop a version of the argument from causation which does not depend on the Phenomenal Principle and the notion of an immediate object of awareness. The argument proceeds by setting up two "job descriptions"[6]: for representations in (A) representation-friendly direct realism, and (B) an explanation of perceptually guided behaviour, and demonstrating that the representations posited by cognitive science, though well suited to role B, are incapable of meeting the requirements of role A.

Starting with the simulation models widely used in science and engineering and then the representations posited in orthodox (Marr, 1982) and connectionist (Churchland, 2002) theories of visual perception, we find that in all cases the representations are second-order isomorphic (Shepard & Chipman, 1970) to their targets and, as such, cannot meet the requirements of role (A). Specifically they cannot supply representational content which represents objects as having the properties that they (the objects) actually have. We conclude that the representations which are, or could be, posited by contemporary cognitive science are incompatible with the metaphysical directness-claim.

However, it could be maintained that Gibsonian representation-averse cognitive science, which rejects the standard causal picture and thereby avoids the causal arguments, has not been shown to be incompatible with the metaphysical directness-claim and thus constitutes a counter example to our central thesis. Drawing on Chapter 3, we would reject this suggestion on two

---

[6] We take the idea of a "job description" from Ramsey (2007).

counts. Firstly, there are, as yet, no fleshed-out representation-averse explanations of the perceptual process, and arguably none in prospect. Secondly, although Gibsonian cognitive science might be compatible with the directness-claim, it is not robustly realist; indeed some of its proponents are avowedly antirealist. Although, outside the scope of this dissertation, we would go further to argue that representation-averse cognitive science is, in principle, incompatible with the realism-claim. We suggest, as the basis for further work, such an argument in Appendix A.

Terminology

It is important, at the outset, to clarify two sources of terminological confusion: firstly, terms such as "naturalism" where, although usage is very broadly consistent, they are open to a range of interpretations and need to be defined more precisely for the purposes of this dissertation; secondly terms, such as "representationalism" and its cognates which are used in conflicting senses.

Two terms of the first type, "metaphysical direct realism" and "naturalistically understood", appear in our central thesis and are thus of particular importance. Since we will argue that much of the literature is compromised by a failure to disambiguate the thesis of "direct realism" which is being defended, clarity on this point is essential. To this end we explicate "metaphysical direct realism" as the conjunction of a metaphysical directness-claim and a realism-claim. (The definition of the term "immediate object of awareness" and its relationship to the notion of phenomenal character, which is crucial to the explication of the metaphysical directness-claim, is also addressed in this section.) Moreover, to avoid a deflationary construal of the thesis at issue, we draw the distinction between a *factual* and *strictly-conceptual* thesis, where "*strictly-conceptual*" is used in the particular sense of placing a principled restriction on the type of data that is allowed to count for or against that thesis. These distinctions are addressed in Section 1.2.

"Naturalism" and its cognates will be less of a concern, since we a adopt a broad-church understanding of the term, essentially the "Muggle constraint", "no spooky stuff allowed"

(Wheeler, 2005, p. 5), and grant that all the theories of perception so far developed within cognitive science constitute naturalistic explanations of the perceptual process. Theories which depart significantly from contemporary physics, say by positing non-local causation at the macroscopic level (Turvey, 2003), and are in this sense not naturalistic, have been suggested but none have been developed in any detail. "Naturalism" is discussed further in Section 1.3.2.

We use the term "argument from causation" to refer to arguments against direct realism which draw on the understanding of the physical and physiological mechanism of perception which came to the fore in the seventeenth century and has been fleshed out by mainstream cognitive science; this we have dubbed the "standard causal picture". Importantly, we grant the possibility of alternative naturalistic non-standard causal pictures which depart from the standard causal picture, say, by espousing extended cognition in the strong constitutive sense, or non-local causation at the level of brain processes. Standard and non-standard causal pictures are discussed in Section 1.3.4.

The term "representationalist", and its cognates, can be extremely confusing, since in the context of perception it is used in two very different senses: representationalist direct realism (Aydede, 2013)[7] (Crane & French, 2015) and Lockean representationalist indirect realism (Armstrong, 1961) (Bonjour, 2013). We will use the term exclusively in the first sense, taking the terms "representationalist" and "intentionalist" to be synonymous; where appropriate we will use the term "representationalist/intentionalist" to emphasize the point. Many versions of representationalism have been discussed in the recent literature, but we will be particularly concerned with "Strong Representationalism" which asserts that "phenomenal character is one and the same as representational content". (Tye, 2009) See Section 4.1.

---

[7] As Aydede puts it, "strong representationalism is the modern day's direct realism" (2013).

The term "naïve realism" is also a source of confusion, since it is used as a synonym for "direct realism" (Dretske, 1995), as a thoroughgoing version thereof which is associated with disjunctivism (Martin, 2004) or to distinguish between "naïve" and "scientific" versions (Dancy, 1985, p. 147). Problems arise when it is unclear whether an objection to "naïve realism" is to be understood in the particular or more general sense. A similar confusion arises with the term "realism", which is sometimes used, particularly in cognitive science (Gibson, 1967), as shorthand for "direct realism" rather than the realism-claim which is shared by both direct and indirect realism. We will need to be mindful on both counts.

The term "causal theory of perception" is used in two different ways. In the older literature, eg (Locke, 1967), it is used as a synonym for indirect realism. We will avoid this usage. However, the term "The Causal Theory of Perception (The CTP)", eg (Grice, 1961) (Vision, 1997), is used to refer to the thesis that, loosely put, the concept of perception necessarily involves causation. Proponents of this thesis are not necessarily committed to indirect realism, in fact Vision is a direct realist. We will use The Causal Theory of Perception as an exemplar when we introduce the notion of a *strictly-conceptual* thesis in Section 1.2.3. However, to avoid confusion, it is important to note, that we have no other interest in the theory and its truth or falsehood is not at issue in this dissertation.

Regarding footnotes

The argument presented in this dissertation does not depend on the extensive footnotes. However, in view of the broad scope of the work, there are inevitably many points where although supplementary material or further discussion might be useful to some readers, it is inessential and inappropriate to the main text. Where there appears to be no need for elaboration, the footnotes can be ignored. Essential material which would otherwise disrupt the flow of the text is included in parenthesis.

Chapter 1: The argument from causation and its logic

Here we are concerned with the argument from causation from the perspective of the philosophical literature. Section 1.1 introduces the argument in the context of all four traditional arguments against direct realism and shows that, although often dismissed as irrelevant or confused, it does in fact present a particularly serious threat to metaphysical direct realism. That the strength of the argument often goes unrecognized, arises from a failure in much of the literature to distinguish between directness v realism, metaphysical v epistemological and factual v conceptual claims; thus, for example, objections proper to a defence of an epistemological claim are improperly used in defence of a metaphysical claim. These distinctions are explicated in Section 1.2.

Having disambiguated its target as metaphysical direct realism, Section 1.3 goes on to formalize the argument from causation as an argument against the conjunction of the metaphysical directness-claim factually construed and the standard causal picture. It is in this form that the argument is required to defend our central thesis. Section 1.4 summarizes our conclusions.

1.1 The arguments against direct realism and objections thereto

Up until the beginning of the twentieth century the majority of philosophers dismissed the direct realist theory of perception as obviously and seriously flawed, but by the final decades it had become the orthodoxy[8]. Robinson makes the point well:

> The majority of modern philosophers – that is, the majority of philosophers writing since the seventeenth century – have believed that in perception one is aware of some item other than the physical object one takes oneself to be perceiving. …On the other hand the majority of strictly contemporary philosophers – that is, the majority of philosophers active in the analytic tradition since the Second World War - ... affirm that we are, normally, directly aware of the external world itself. (1994, p. 1)

More recently Fish writes:

---

[8] There are a small number of exceptions, including Robinson (1994), also Jackson (1977) and Foster (2000).

> There was a time when to call a theory of perception a version of "direct realism" was almost equivalent to calling it "hopelessly naïve". Time has told however, that it was this assumption which was naïve, and nowadays the majority of theories on the market see themselves as direct realist. (2004)[9] [10]

Less well recognized perhaps, is that this shift towards direct realism arose not out of a period of intense debate, but following a period of relative inactivity. Writing in the mid sixties Warnock notes the lack of material for his contribution to the *Oxford Readings in Philosophy*:

> … a collection of recent or contemporary articles on perception would have been easier to make some twenty or more years ago than it is today. There would have been more to choose from ….[I believe that] philosophical interest in perception is at present rather in a state of suspended animation. (1967, p. 1)

More recently, Maund writes:

> … for a lengthy period in the twentieth century, the topic of perception and its central issues were largely ignored. What is additionally remarkable is that for most of the first half of that century the topic was of paramount significance. (2003, p. x)

Loosely put, philosophers started the twentieth century as indirect realists or phenomenalists, lost interest during the middle years, and espoused direct realism on their return to the topic in the last few decades. Indeed Martin talks in terms of "a discontinuity in the debate." (2000, p. 4)

In contrast, the majority of cognitive scientists have remained, roughly speaking, indirect realists. Although there is a significant Gibsonian heterodoxy, mainstream cognitive science is still Helmholtzian and rejects the notion of direct perception. (Noë, 2002) (Smythies, 2009) The debate within cognitive science is discussed in detail in Chapter 3. Here, and in Chapter 2, we

---

[9] According to Dancy: "For the period 1900 – 50 most philosophers were indirect realist. Now the pendulum has swung in favour of direct realism." (1985, p. 159) Vision talks of a "near consensus direct realism." (1997) Le Morvan talks of "the demise of the Sense-Datum Theory and Phenomenalism in the last century." (2004) Crane confirms the shift from sense data theories, indirect realism and phenomenalism, as serious options: "…it seems fair to say that while at the beginning of the 20th century, the chief protagonists in the philosophy of perception were direct realists and the sense data theorists, at the beginning of 21st century they were disjunctivists and intentionalists." (2005, p. 261)

[10] But note Matthen's somewhat lighthearted comment in the 2010 PhilPapers discussion: "DR is not the majority view in philosophy (though realism is). (Jonathan, the fact that both Snowdon and Martin are at UCL may have given you the impression that most philosophers are in this camp.) As a matter of pure description, most direct realists live in England or were born there." http://philpapers.org/bbs/thread.pl?tId=552 However it appears that by "DR" Matthen is referring to "naïve realism" (Martin, 2004), that is disjunctivism rather than representationalism/intentionalism.

focus on the philosophical literature. However, in advance of the discussion below, it is important to emphasize that although direct perception may be the orthodoxy in contemporary analytic philosophy, it remains the heterodoxy in cognitive science. Moreover, although outside the scope of this dissertation, there is an implicit acceptance of a loosely formulated indirect realism amongst much of the wider scientific community.

The implicit acceptance of, or at least sympathy towards, indirect realism amongst the scientific community is indeed widespread as exampled by five very different sources. From *The Astonishing Hypothesis*: "We have for example a vivid internal picture of the external world." (Crick, 1994, p. 9)[11] From the *God Delusion*: "What we see of the real world is not the unvarnished real world but a *model* of the real world, regulated and adjusted by sense data – a model that is constructed so that it is useful for dealing with the real world." (Dawkins, 2006, p. 371) From an IEEE symposium on virtual reality: "I maintain that understanding synthetic experience must begin with recognizing the fallacy of naïve realism and with the recognition that the phenomenology of synthetic experience is continuous with that of ordinary experience." (Loomis, 1993, p. 54) And from the perspective of physics and cosmology in *The Grand Design*: "According to the idea of model-dependent realism ..., our brains interpret the input from our sensory organs by making a model of the outside world. …It follows that a well-constructed model creates a reality of its own." (Hawking, 2010, p. 217) Decades earlier, Einstein is taken to have "expressed great admiration" for Russell's defence of indirect realism. (McLendon, 1956)

---

[11] Exemplifying the tension with the direct realism of contemporary analytic philosophy, Searle takes issue with Crick: "He is even led to deny that we have direct perceptual awareness of objects in the world" and on this count he has been "badly advised philosophically". (1998, p. 32)

1.1.1 The traditional arguments against direct realism

There were very few direct realists until the middle of the last century[12]. Reid, a lone voice writing towards the end of the eighteenth century, remarks that "All philosophers, from Plato to Mr Hume, agree on this …. That we do not perceive external objects immediately, and that the immediate object of perception must be some image present to the mind." Indeed, "there appears to be a unanimity rarely to be found among philosophers on such abstruse points." (Reid, 1785, p. 117) The reasons for the near unanimity can be found in the general acceptance of objections to the notion of direct perception which later came to be known as the arguments from illusion, hallucination, causation and science.

Precursors of all four arguments can be found in the Greeks and the Early Moderns, and the arguments proper were discussed more rigorously in the early twentieth century. However there has been a shift of emphasis, broadly from scientific considerations in the seventeenth century to the problems associated with illusion and hallucination in the early twentieth. The roots of the argument from illusion/hallucination can be traced back to the Sceptics (Hamlyn, 1961, 32), and it is strongly suggested in the Early Modern period in Descartes, Locke, Berkley and Hume. However the argument in its current form only came into prominence early in the twentieth century with the development of the notion of sense data.[13] According to Robinson it "figured hardly at all in the writings of the early modern philosophers" (1994, p. 59), and it was in the early twentieth century that the phenomena of illusion "tended to replace scientific facts as the basis for arguments in the philosophy of perception." (1994, p. 37) On the other hand the arguments from science and from causation, the precursors of which date back to the Atomists

---

[12] Searle talks of the "embarrassing" fact that "if you look at the history of philosophy from Descartes on, there were no naïve realists among the great philosophers." (2003)

[13] Against McDowell et al, who interpret Descartes reflections in the First Meditation as a form of the argument from illusion, McArthur argues that they are better understood as "offering a causal argument for scepticism that depends on a naturalistic conception of sense experience." (Macarthur, 2003, p. 175)

(Hamlyn, 1961, p. 8), were at their most influential with the scientific advances of the seventeenth century and, as Robinson suggests, had declined in influence amongst philosophers by the twentieth.

Although nothing rests on this historical sketch, the reliance on the arguments from illusion and hallucination by phenomenalists and indirect realists in the early twentieth century might go some way to explaining why, some decades later, direct realists directed their fire at these arguments rather than arguments which addressed the process of veridical perception. Significantly, Austin's Sense *and Sensibilia* (1962), which was influential in the shift to direct realism, focuses almost exclusively on the arguments from illusion and hallucination and has little if anything to say about the process of veridical perception. Moreover, mid twentieth century Linguistic Philosophy[14] encouraged the view that considerations relating to the physics and physiology of perception, on which the argument from causation depends, are somehow irrelevant to the "philosophy of perception". The vestiges of Linguistic Philosophy remain and such a demarcation is still implicit in some recent work. However, as we will see in Chapter 3, the claim that there is a "philosophy of perception" which can be divorced from a "science of perception" is difficult to maintain in the context of recent philosophically sophisticated cognitive science (call it scientifically savvy philosophy if you wish) as exemplified by the Churchlands (2002) (2012), Wheeler (2005) and Chemero (2009).

1.1.2 Twentieth century objections

So how did these four arguments, the precursors of which had held sway since the Early Modern era suddenly, over just a few decades in the middle of the twentieth century, fall into disrepute and direct realism become the orthodoxy amongst analytic philosophers? To address this question we need to understand the logical structure of the arguments.

---

[14] As exemplified by the philosophy of the later Wittgenstein (1953), Ryle (1953), Austin (1962).

We begin with Snowdon's distinction between the "base case" and "spreading step". (1992) With respect to the arguments from illusion and hallucination, the "base case" makes the claim that in a particular illusory or hallucinatory case we do not directly perceive a physical object, whilst the "spreading step" extends the scope of this negative conclusion to the veridical case.

Consider the argument from illusion. In the example of the partially immersed stick, the base case seeks to establish that, in the illusory case where the stick appears bent, the immediate object of awareness cannot be one-and-the-same-as the physical object, since the one is bent and the other is straight. The spreading step then seeks to extend this conclusion to veridical perception, to claim that the immediate object of awareness is never the physical object, even when no illusion is involved.

However, as the late twentieth century critics of the arguments pointed out, the base case and spreading step depend, respectively, on premises which came to be known as the "Phenomenal Principle" and the "Common Factor Principle":

> [Phenomenal Principle] If there sensibly appears to a subject to be something which possesses a particular sensible quality then there is something of which the subject is aware which does possess that quality. (Fish, 2010, 6)
>
> [Common Factor Principle] Phenomenologically indiscriminable perceptions, hallucinations, and illusions have an underlying mental state in common. (Fish, 2010, 4)[15]

Although often implicitly assumed, but as Smith makes explicit (2002, p. 8), the arguments also draw on the principle that discernibles cannot be identicals; that is, if X and Y differ in any respect they cannot be one and the same. We dub this the "Indiscernibility Principle". It will be significant with respect to the argument from causation and is for this reason made explicit here in our discussion of the general structure of the arguments.

---

[15] The two principles were introduced by Robinson (1994) and Martin (2004) respectively. Though Martin uses the term "Common Kind Assumption".

To see the three principles in operation, consider again the argument from illusion and the partially immersed stick. The Phenomenal Principle is needed to sanction our use of the notion of an immediate object of awareness. It is only by calling on the Phenomenal Principle that we can justify the claim that there is something, an immediate object of awareness, which has the quality of being bent. The Indiscernibility Principle is required to support the claim that because the immediate object of awareness and the physical object are discernible (one is bent and the other is straight) they are not identical.

The Common Factor Principle is needed to justify generalizing the conclusion of the base case to the case of veridical perception. Unless the Common Factor Principle is accepted, the argument is open to the objection that although the immediate object of awareness and the physical object are not identical in the illusory case, they are nonetheless identical in the veridical case. Indeed the direct realist might accept that, although the immediate object of awareness is a physical object in the veridical case it is a sense-datum in the illusory case.

From the 1950s on, philosophers developed three versions of direct realism, "The Intentionalist Theory" (aka representationalist theory) "The Adverbial Theory", "and "The Disjunctivist Theory" (Crane, 2011) (Crane & French, 2015)[16] which arose out of three different objections to the arguments from illusion and hallucination. The intentionalist and adverbialist objections are directed at the base case: they reject the Phenomenal Principle and refuse to sanction the notion of an immediate object of awareness. The disjunctivist objection is directed at the spreading step; it rejects the Common Factor Principle by drawing a radical distinction between the objects of veridical and non-veridical perception.

---

[16] Interestingly, although Crane uses the headings "3.1 The Sense-Datum Theory", "3.2 The Adverbial Theory", "3.3 The Intentionalist Theory" and "3.4 The Disjunctivist Theory" in the Spring 2011 edition of the Stanford Encyclopedia, he uses "3.4 The Naïve Realist Theory" in the Fall 2015 edition.

The intentionalists and disjunctivist objections are the big guns of the late twentieth century rehabilitation of direct realism.[17] Significantly, the battles were fought over illusion and hallucination, rather than veridical perception and the argument from causation which received surprisingly little attention.

Whether or not the intentionalist and disjunctivist objections do in fact dispose of the arguments from illusion and hallucination, we can leave as an open question, though there are reasons to suppose that these objections are at best inconclusive.[18] Indeed, there is reason to suppose that the forces opposed to direct realism retreated in the face of a changing intellectual climate, rather than under the weight of fire from objections to the arguments from illusion and hallucination.[19] However we can put these issues aside; the argument from causation stands or falls independently[20].

---

[17] Although adverbialism was the first, and for some time the most influential, of the trio of objections it now has few advocates and will not be discussed further. Jackson's "many properties" objection (Jackson, 1977, p. 63) (Robinson, 1994, p. 180) has been particularly influential and is now widely accepted. Nothing depends on the exclusion of the adverbialist theory from the debate.

[18] The representationalist and disjunctivist counters to the traditional arguments from illusion and hallucination are at best inconclusive, since representationalism and disjunctivism are themselves controversial and conflicting theories. Consider the responses to the argument from hallucination. Armstrong (1961), Huemer (2001) and Smith (2002) all block the argument at the base case with a representationalist objection to the Phenomenal Principle, but reject disjunctivism and endorse the Common Kind Principle. On the other hand Martin (2004) McDowell (1982) and Brewer (2007) all endorse the base case and block the argument with the disjunctivist objection to the Common Kind Principle at the spreading step. The one position seriously undermines the other. Advocates for the argument from hallucination can side with Smith et al with respect to their endorsement of the Common Kind Principle and McDowell et al for an endorsement of the Phenomenal Principle. (Griffiths, 2008) The continuing dispute within the direct realist camp is well illustrated by Searle's recent no holds barred attack on disjunctivism. (2015, p. 164)

[19] Even if we were to ignore these disagreements within the direct realist camp and give the benefit of the doubt to the late twentieth century objections to the arguments from illusion and hallucination, they (the objections) are not nearly strong enough to explain direct realism's dramatic rise from hopelessly naïve to the late twentieth century orthodoxy. The fine distinctions that need to be drawn, the slippery nature of the concepts involved and complexity of the argumentation are such that these deliberations are unlikely to be convincing. It seems that there must be other factors at play. Indeed Martin explains the dramatic shift in the philosophy of perception in terms of the changes in intellectual climate which took place in the middle of the twentieth century. He argues that the "remarkable shift" from sense data theories to intentional theories arose not out of developments in the substance of the debate, but out of changes in the "intellectual context" in which the debate was conducted and specifically "the ascendancy of a commitment to physicalism". (Martin, 2000, p. 32)

[20] Though as discussed in Chapter 2, Armstrong maintains that the argument from causation draws on the argument from illusion. We disagree.

Indeed, in comparison to the arguments from illusion and hallucination, the argument from causation has two particular strengths. Firstly, since the argument from causation addresses veridical perception from the outset, it does not involve the spreading step, does not draw on the Common Factor Principle, and thus does not face the disjunctivist's objection.

Secondly, the representationalist's objection to the Phenomenal Principle is blunted when only the veridical case is at issue. It is in Robinson's original form (1994, p. 32) that the Phenomenal Principle is required by the arguments from illusion and hallucination, and it is at this form that the representationalist objection is usually directed. However since illusory or hallucinatory cases are not at issue, the argument from causation requires only a weaker and less contentious version of the principle. We have dubbed this the "Weak Phenomenal Principle":

> *In the case of veridical perception* [italicized words added to Robinson's original]*, if there sensibly appears to a subject to be something which possesses a particular sensible quality then there is something of which the subject is aware which does possess that sensible quality. (Robinson, 1994, p. 32)

The Weak Phenomenal Principal does not claim that there is an immediate object of awareness in the case of illusion or hallucination. Consequently, compared to that required to counter the arguments from illusion and hallucination, a more thoroughgoing form of representationalism is required to counter the argument from causation. To do so, it must be denied that there is an immediate object of awareness in the case of veridical perception. Starkly put, with no ameliorating reference to illusion or hallucination, such a denial appears to be at odds with the spirit of direct realism. Smith characterizes this denial as the "feigned anaesthesia" of early behaviourists. (2002, p. 46)

However, although these considerations raise serious issues with respect to the plausibility of the representationalist objection, they do not dispose of it. Although rarely stated explicitly, the rejection of the Weak Phenomenal Principle is implicit in Armstrong's analysis of perception as "nothing but" the acquisition of beliefs (1961) and in recent reductive representationalist theories

which identify the phenomenal character of experience with representational content. (Tye, 2009) These issues are addressed in Chapter 2 in the context of Armstrong's objections to the argument from causation and, in Chapter 4, motivate a detailed examination of the representations posited within mainstream cognitive science.

Thus the argument from causation spikes the two big guns of the late twentieth century defence of direct realism; the disjunctivists' objection is silenced and the representationalists' objection is, at least partially, disarmed. In this respect the argument is robust. Nevertheless it is often dismissed lightly as confused, on the grounds that considerations with respect to the physical and physiological mechanism of perception are irrelevant to the standing of direct realism and, consequently its (the argument's) particular strengths go unrecognized. We dub this the "irrelevance objection".

However we will maintain that the charge of irrelevance arises from a failure to disambiguate the thesis of "direct realism" which is being defended and, moreover, that the irrelevance objection has no purchase against an argument which is clearly targeted on metaphysical direct realism's directness-claim factually construed. Such an argument is presented in Section 1.3 and defended in Chapter 2, but first we need to establish four key distinctions which are essential to disambiguating the thesis which is at issue in this dissertation.

1.2 Four key distinctions

Here we introduce four key distinctions: between directness and realism claims, between metaphysical and epistemological claims, between factual and conceptual claims, and between arguments against direct realism and argument for some alternative theory of perception. As we will see, much of the literature relating to the argument from causation is compromised by a failure to observe these distinctions.

The term "immediate object of awareness" plays a key role in drawing these distinctions. We follow Smith (2002) in using this term to refer to that of which we are immediately aware, in a sense which is neutral between direct realism, indirect realism and phenomenalism; on this understanding, the question at issue is the relationship between the immediate object of awareness and the physical object.[21] This is the familiar sense in which Armstrong uses the term in the opening paragraph of his seminal monograph *Perception and the Physical World*:

> There is a triad of 'theories of perception' which compete for the allegiance of philosophers: Direct Realism, Representationalism[22] and Phenomenalism. Each may be conceived of as an answer to the question 'What is the *direct* or *immediate* [italics in original] object of awareness when we perceive?' (1961, p. xi)[23]

At this point it is also useful to clarify our understanding of the relationship between the immediate object of awareness and phenomenal character. On our understanding, we take phenomenal character to be a property of the immediate object of awareness. Indeed we might introduce the term "immediate object of awareness" phenomenologically, as that which bears the "phenomenal character" which is presented in perceptual experience.

However, we recognize that, as Fish (2009) points out, the literature divides on this matter. Some (eg Fish cites Byrne, Chalmers and Levine) maintain that the phenomenal character of an experience, when properly understood, is "a property of the experience itself." On the contrary, others (eg Martin and Dretske) maintain that "the features which define what it is like to have an experience ", that is its phenomenal character, "are properties the objects we experience (not our experience of them) have." (Fish, 2009, p. 10) As to whether perceptual character is properly understood as primarily a property of experiences or of objects, we take the latter position.

---

[21] As Huemer notes, it is in this sense in which the term "sense data" was originally used when it was introduced in the early twentieth century "when it was supposed to be neutral between direct and indirect realist theories of perception, so that it was not to be assumed either that sense data must by definition be mind-dependent or that they must be mind-independent." (2011) The term ("sense data") is of course used differently in the more recent literature.

[22] Note that Armstrong is using the term "Representationalism" and its cognates to refer to indirect realism not representationalism/intentionalism.

[23] As we noted in Section 1.1.1, Reid uses a similar term when he takes issue with the claim that that "the immediate object of perception must be some image present to the mind." (Reid, 1785, p. 117)

1.2.1 Directness v realism claims

Direct realism makes two claims: the directness-claim and the realism-claim. As Smith puts it "Direct Realism" is the "position that combines the issue of directness with a Realism about the physical world." (2002, p. 1) The realism-claim is that which unites direct and indirect realism against phenomenalism, and the directness-claim is that on which the two realisms divide.

Metaphysical and epistemological direct realism make the same "realism-claim":

> The physical world has an existence that is not in any way dependent upon it being "cognized". ….the purely physical elements of this physical world of ours would, or at least could, be just as they are even if no perceiving or thinking subjects had ever come into existence. (Smith, 2002, p. 2)

With the demise of phenomenalism, the realism-claim is rarely questioned in the contemporary analytic philosophy of perception and often passes unnoticed in a debate which focuses on the directness-claim. Indeed, the realism-claim appears to be an unspoken assumption in both the (analytic) philosophy and (mainstream) cognitive science of perception and will only come to the fore in this dissertation when we address Gibsonian cognitive science in Chapter 3.[24] Significantly, parallels have been drawn between Gibsonian cognitive science and the Phenomenological Tradition in which the likes of Merleau-Ponty, Husserl and Heidegger have been taken to question the realism-claim.[25]

 It might be objected that a direct realist can be true to the cause without subscribing to the realism-claim as we have presented it above. We address this issue in Chapter 3, where we

---

[24] Although outside the scope of this dissertation, there are parallels with the ongoing debates about realism in the interpretation of quantum mechanics as, for example, regarding the EPR debate (Einstein, et al., 1935), "The Quantum Theory and Reality" (d'Espagnet, 1979) and "Realism and Objectivism in Quantum Mechanics" (Karakostas, 2012). It is significant that non-realist interpretations are given serious attention and are not taken to be in opposition to scientific naturalism.

[25] Smith (2002, p. 4), for example, in discussing their treatment of the problem of perception, maintains that Merleau-Ponty and Husserl are idealists, though he accepts that the issue is contentious. He takes Heidegger's position to be more ambiguous. On this see *Heidegger: Between Idealism and Realism* (Stephanich, 1991), and *Heidegger's Reading of Descartes' Dualism: The Relation of Subject and Object* (Cucen, 1998). We will return to these issues in Appendix A in the context of the realism-claim, Merleau-Ponty's meaning laden world and Heidegger's dissolution of the Cartesian subject/object dichotomy.

suggest that Gibsonian radical embodied cognitive science is at odds with the realism-claim as it is actually maintained by contemporary direct realist analytic philosophers. (An argument to his effect is proposed in Appendix A.) Up until that point, nothing rests on the detail of the realism-claim. It is sufficient to have established the distinction between the realism-claim and the two directness claims.

1.2.2 Metaphysical v epistemological claims

Two types of issue arise in the philosophy of perception, metaphysical and epistemological, which give rise to two questions: (i) What is the direct or immediate object of awareness when we perceive?, and (ii) How can perception give us knowledge of the external world? [26] Depending on which question is being addressed, "direct realism" can be presented as a theory about either the metaphysics or epistemology of perception. It is with respect to the directness-claim that metaphysical direct realism and epistemological direct realism divide.

Using Smith's form of words[27], we take the metaphysical directness-claim to be that:

> … the immediate object of awareness in standard perceptual situations is a normal physical object - in other words, that it is *identical* to some such object. (Smith, 2002, p. 7)

Importantly, we follow Smith in presenting the metaphysical directness-claim as a positive claim about what the immediate object of awareness is, that is a normal physical object, rather than a purely negative claim about what it is not, that is a perceptual intermediary of some sort.[28] [29]

---

[26] See the two Stanford Encyclopedia entries on the problem of perception: *The Problem of Perception* (Crane & French, 2015) and *Epistemological Problems of Perception* (Bonjour, 2013).

[27] But see the caveat below regarding Smith's failure to distinguish between metaphysical and epistemological direct realism.

[28] Smith uses "immediate object of awareness", or similar locutions, as we use the term, throughout his treatment of the arguments. Regarding the argument from illusion, he proposes a "subtle and non-obvious" (2002, p. 22) response which draws on the phenomenology of perceptual constancy and the distinction between the changing "[perceptual] sensation" and the unchanging "immediate object of perceptual awareness." (p. 66) Regarding the argument from hallucination, he rejects the "denial that a hallucinating subject is aware of anything" (p. 195), rather the "object(s) of awareness" are ""nonentities" or "non-existents""(p. 234); though he recognises that this is a controversial notion and devotes the last 35 pages to its defence.

Armstrong, like Smith, is explicit on this point and introduces direct realism as the theory of perception which maintains that "the immediate object of awareness is never anything but a physical existent." (1961, p. xi) More recently Brewer defends the claim that "these direct objects [presented in experience] are the persisting mind-independent objects we all know and love." (2007, p. 87) Where direct realism is presented as the negative claim, as in Huemer's "having awareness of [a thing] not based on one's awareness of anything else" (2001, p. 135), as discussed in Section 2.2.2, we take the corresponding positive claim to be implicit. Indeed we reject a construal of metaphysical direct realism, and a defence thereof, which maintains only the negative claim as deflationary and inadequate.

Although, in defending our central thesis, we are concerned with metaphysical direct realism, and thus the metaphysical directness-claim, we need to present the epistemological directness-claim for the purposes of differentiation and comparison. Nothing depends on the precise formulation of the epistemological directness-claim. The key point is that it concerns the basis of perceptually acquired knowledge, or belief, and denies that such is based on inference.

Drawing on Fumerton, we take the epistemological directness-claim to be that:

> … we have noninferentially justified beliefs in at least some contingent propositions describing the external physical world. (2006, p. 680)

However, the literature often fails to draw the distinction between metaphysical and epistemological direct realism and it is unclear which version of the thesis is being defended.

---

[29] Since we follow Smith, in both his definition of direct realism as an identity claim which draws on the notion of an "immediate object of awareness" and his understanding of the structure of the arguments against (Section 1.3.2), the question arises as to his rationale for defending the very thesis that we reject. Two points in response. Firstly he is concerned exclusively with the arguments from illusion and hallucination and suggests that direct realism might be rejected on other grounds (see fn 70); we might take this as a nod in the direction of the argument from causation. Secondly, although he does defend direct realism defined in our terms, he recognises the difficulty in so doing and the inadequacy of the standard objections to arguments from illusion and hallucination; his monograph is advertised as answering arguments against direct realism that "no current theory of perception can adequately rebut." (Smith, 2002) Indeed, we might argue that although he eventually finds in favour of direct realism, he does more harm to the cause by undermining the arguments of other direct realists than he benefits the cause by presenting his own "subtle and non-obvious" (2002, p. 21) arguments.

Even Smith's impressively detailed monograph which defends direct realism against the arguments from illusion and hallucination fails to draw the distinction. Indeed, although we draw on Smith's form of words to express the metaphysical directness-claim, he does not label it as such; rather, it is the "identity claim" (no reference to metaphysical or epistemological) which "Direct Realism must be understood as making". (Smith, 2002, p. 8)

As we will see, despite warnings to the contrary (Dretske, 2004) (Thompson, 2008), failure to draw this distinction is still widespread in the literature defending direct realism. A notable example is le Morvan's *Arguments Against Direct Realism and How to Counter Them* (2004) discussed in Section 2.2.2. However, the distinction is crucially important, since metaphysical and epistemological direct realism stand or fall independently[30], and sound arguments against the one may be unsound against the other. Indeed, Thompson (2008) suggests that Smith (2002) is, in effect, rejecting metaphysical direct realism and only defending epistemological direct realism. Considering the depth and thoroughness of Smith's monograph (2002), that there could be such a fundamental ambiguity in the thesis which he is defending presents a salutary warning.

1.2.3 Factual v conceptual claims

Since at least the Early Moderns (eg Reid, Locke and Berkeley) up until the middle of the last century, the three theories of perception were assumed to be making important factual claims about this world in which we find ourselves, much as the Copernican and Ptolemaic theories made such factual claims. This understanding underpins and motivates the arguments presented in this dissertation.

---

[30] For example, it could be argued, contrary to the epistemological directness-claim, that unconscious inferences are involved in making perceptually guided judgments, say that object A is nearer than object B, even if in accordance with the metaphysical directness-claim it is accepted that the immediate object of awareness is one-and-the-same-as the physical object. On the other hand, it could be argued, perhaps on the basis of a "reliabilist" theory of knowledge (Goldman, 2015), that, in accordance with the epistemological directness-claim, inference is not involved, even if the metaphysical directness-claim is rejected in favour of a sense data theory.

However, although this construal of the theories of perception carries through into contemporary cognitive science and much of the philosophy of perception, there are other understandings, not always made explicit, on the basis of which direct realism is defended as, in some sense, a "conceptual thesis" which stands independently of facts and theories concerning the physical and physiological mechanism of human perception. Since it will be important to distinguish between a *factual* and a *strictly-conceptual* construal of the direct realist thesis in our discussion of the argument from causation, we need to make clear the basis of this distinction and to define our terms. We use the term *strictly-conceptual* to emphasize that such a thesis is "strict" in the sense that it places a restriction on the type of data that is allowed to count for or against that thesis. (We will italicize throughout to emphasize that these terms are being used in a particular sense.) Although the restriction is variously applied, it is generally taken to disallow empirical scientific data. On the other hand, a *factual* thesis imposes no such restrictions. Importantly, a *factual* thesis does not rule out the data which is ruled in by a *strictly-conceptual* thesis. A *factual* thesis takes on all-comers.

The suggestion that the problem of perception should be construed as a debate between *strictly-conceptual* theses, which can be addressed without reference to the perceptual science of the day, only came to the fore with the Linguistic Philosophy of the mid twentieth century. For Ryle, writing from the perspective of Ordinary Language Philosophy, the "problems of perception" are *strictly-conceptual* in the sense that they are properly concerned with "how certain words work":

> If we are enquiring into problems of perception, i.e., discussing questions about the concepts of seeing, hearing and smelling, we may be taken to be tackling the questions of opticians, neuro-physiologists or psychologists, and even fall into this mistake ourselves. It is then salutary to keep on reminding ourselves and one another that what we are after is an account of how certain words work, namely words like 'see,'look', 'overlook', 'blind', 'visualise' and lots of other affiliated expressions." (Ryle, 1953, p. 185)

Few philosophers with a serious interest in perception would now be willing to accept that they were merely concerned with "how certain words work"; indeed, even in the 1950s, Ryle clearly

recognizes that some of his fellow philosophers would be tempted to depart from this narrow conception of their role and needed to be reminded thereof.[31] However the assumption that the "problems of perception" concerns, exclusively, questions about "concepts" and, furthermore, that such questions can be addressed without reference to the findings of the "neuro-physiologists or psychologists" is sometimes implicit in the more recent literature.

Fumerton is explicit on the matter and argues that "science has no light to shed on either the epistemological or metaphysical problems of perception." (1985, p. 7) He explains why the "scientist's attempts to characterize the neurophysiological workings of perception ...fail to answer the metaphysical questions of perception":

> Light may impinge on the retina and by a complex process result in an alpha wave in the brain, but this is no part of what we assert, no part of what we *mean* when we say we see a table. (Fumerton, 1985, p. 23)

Significantly he has to go to some lengths to explicate and defend his proposed method of "meaning analysis" and to distinguish it from the "intellectually less lofty" activities of the lexicographer. (p. 14) As he puts it: "It is perhaps an understatement to suggest that the conception of philosophical analysis presented above is not all that popular today." (p. 23)

The clearest example of a current debate where a *strictly-conceptual* thesis is at issue is with respect to "The Causal Theory of Perception". This debate was initiated by Grice (1961) and continues through to the present. Here the issue is whether the concept of perception necessarily involves the notion of causation; proponents of the theory maintain that it does. For our purposes, the key issue is the way in which The Causal Theory of Perception places a restriction on the admissible data. In his seminal paper Grice rules out reference to the "specialist's account" in terms of "the transmission of light waves to the retina" on the grounds that:

---

[31] Significantly, this narrow conception of the proper role of philosophy was associated with a metaphysical assumption that, as Gellner puts it in his then controversial *Words and Things*, "The World is as it seems." (1959, p. 124)

> …if we are attempting to characterize the ordinary notion of perceiving, we should not explicitly introduce material of which someone who is perfectly capable of employing the ordinary notion might be ignorant. (Grice, 1961, 143)

Vision (1997) defends the theory in a substantial monograph where he presents it as a "conceptual thesis" (p. 11) which restricts the data to that available to the "many lay possessors" of the concept of perception (p. 13). Snowdon examines this type of theory and, more recently endorsed by Fish (2010, p. 118), suggests that a distinctive feature of the "conceptual truth" which it asserts is that "there is a restriction on the data to which appeal can be made in the supporting argument." Specifically:

> … the data must be relatively immediately acknowledgeable by any person, whatever their education, who can count as having the concept in question. The aim of the restriction is to exclude any facts of which we can become aware only in the context of some activities (for example, carrying out experiments, or becoming acquainted with the results of experiments, or reading psychological text books) which need not be indulged in by just anyone who has the concept. (Snowdon, 2002, p. 152)

So, here, in the context of the philosophy of perception, we have three examples of *strictly-conceptual* theses which, respectively, restrict admissible data to (i) that concerned with "how certain words work", (ii) that concerned with "meaning analysis" and (iii) that which is "acknowledgeable by any person, whatever their education …" On the other hand, a *factual* thesis imposes no such restrictions.

Two points need to be emphasized. Firstly, as should be clear from these examples, our distinction between *strictly-conceptual* and *factual* theses does not rest on the controversial distinction between analytic and synthetic truths. (Quine, 1951) Whilst accepting Quine's dissolution of the latter, a distinction might still be drawn which, as in the case of "The Causal Theory of Perception", rules out reference to the "specialists account" of the perceptual process. (Grice, 1961, p. 143) Secondly, and most importantly, the examination of a putative *factual* thesis must take full regard of "conceptual" considerations, however they are defined. For example, considerations with respect to the lay usage of a concept, may be important

preliminaries, though they are unlikely to be definitive unless the thesis is explicitly concerned with, say, the preconceptions of lay users. Moreover, an analysis of, and perhaps a consequent revision of, expert usage of key concepts may play a crucial role in the development of a *factual* thesis. Indeed, in this respect, some *factual* theses are conceptually heavy. For example, the Special Theory of Relativity is a *factual* thesis which involves a major revision to the Newtonian conception of space and time. (Feynman, 1963)

At this stage we need only establish the distinction and to define the terms *factual* and *strictly-conceptual*. However we will maintain that a *strictly-conceptual* construal of "direct realism", though once sufficient to meet the limited aspirations of Linguistic Philosophy, cannot do justice to the letter and spirit of the thesis which is defended within contemporary analytic philosophy. There is more at stake[32] than such a deflationary construal would allow.

1.2.4 Negative v positive arguments

It is important to distinguish between purely negative arguments against a particular theory of perception, and positive arguments for an alternative theory. This is an important distinction, since we may be capable of refuting a particular theory, whilst being incapable of offering any alternative theory in its place. For example, it is well within our unaided cognitive powers to refute the theory that 2 is a factor of 3733377089, but far beyond to find the true theory as to the factors of this 10 digit number. However our disproof of the theory that 2 is a factor of 3733377089, since the number is odd, is in no way compromised by our inability to arrive at the true theory.

Modesty, with respect to positive claims, may be particularly appropriate in the case of theories of perception. As McGinn points out in the context of the mind-body problem there is "no

---

[32] That is, as Noë puts it, "one's conception of one's place in the natural world. (Noë, 2002, 12) Quoted in full above.

guarantee that our cognitive powers permit the solution of every problem that we recognize." (1997) Moreover, irrespective of any limitations of the cognitive powers of our species, the problem of the relationship between the perceiver and the perceived in a particular world, may be a problem which is, in principle, not resolvable by perceivers in that world. The problem does have a worryingly self-referential form. Nothing rests on this speculation, but it does warn against accepting direct realism as the "least worst" theory on offer.

Although the argument from causation is often used as a positive argument for indirect realism, here it is presented as a purely negative argument against direct realism which does not, and need not, make any positive claims as to a true theory of perception. We will need to draw on this distinction when we counter objections to the effect that the argument from causation has not made the case for some particular indirect realist theory, say, in which brain processes are the immediate objects of awareness. Moreover, as we will see, the conjunctive form of our formalised argument, that the standard causal picture and the metaphysical directness-claim cannot both be true, precludes the making of any such positive claims.

1.3 Refining the argument from causation

Here we refine and formalize our argument against metaphysical direct realism. However, first, it is useful to recognize the resilience of our target and the hold that it has on ones thinking. Strawson puts it well. Before subjecting the "scheme of common sense" to critical examination, he notes that it is essential to stress:

> …the grip that common-sense non-representational realism [direct realism[33]] has on our ordinary thinking. It is a view of the world which so thoroughly permeates our consciousness that even those who are intellectually convinced of its falsity remain subject to its power. [and quoting Mackie approvingly] …even when we are trying to

---

[33] Strawson is clearly using the term "representational realism" to refer to Lockean indirect realism rather than representationalist direct realism; thus we take "common-sense non-representational realism" to be a reference to direct realism though he does not use the latter term. This usage is further evidenced a few pages earlier: "The theory of perception associated with scientific or Lockian realism is commonly and reasonably described as a representative theory." (1979, p. 49)

> entertain a Lockian or scientific realism 'our language and our natural ways of thinking keep pulling us back' to a more primitive view. (Strawson, 1979, p. 53)

Although Strawson does not draw the distinction, it seems clear that the "view of the world which so thoroughly permeates our consciousness" is that of metaphysical rather than epistemological direct realism. Epistemological direct realism does not have such a hold over our thinking. It would not be too difficult to accept, if we were intellectually convinced, that our perceptual judgments depend on inference. On the other hand, to accept that the immediate objects of awareness may not be the "mind-independent physical objects we all know and love" (Brewer, 2007, p. 87), is a very different matter. *Unheimlichkeit*.

1.3.1 The scope of the argument

Drawing on the above four distinctions, we can disambiguate the thesis of "direct realism" which is at issue in this dissertation and against which we are targeting the argument from causation. The thesis at issue is metaphysical direct realism and the argument is targeted on the metaphysical directness-claim *factually* construed. We do not defend the argument as an argument against the realism-claim, the epistemological directness-claim or any *strictly-conceptual* claim. Moreover we do not present the argument as a positive argument for any alternative theory of perception.

The realism-claim is not incompatible with the standard causal picture and an explanation of the perceptual process in terms of a naturalistically understood causal chain of events. Indeed, the realism-claim is implicit in such an explanation and underpins the debate in both mainstream cognitive science and analytic philosophy. The realism-claim will only become an issue when we examine the realist credentials of radical embodied cognitive science in Section 3.3.3.

It might be maintained that essential features of the standard causal picture imply that perception must involve inference and that the standard causal picture is therefore incompatible with the epistemological directness-claim. We leave open the possibility of such an argument; since we

are targeting the metaphysical directness-claim, and an argument against the epistemological directness-claim would take a different form, we confine our energies to the former. However, as discussed in Chapter 3, when the standard causal picture is fleshed out within mainstream cognitive science it does appear to be incompatible with the epistemological directness-claim.

The argument from causation has no purchase against direct realism if it (direct realism) is presented as some type of *strictly-conceptual* thesis. Indeed, in defence of a *strictly-conceptual* thesis, the proponent of direct realism would be entitled to disallow the causal considerations, drawn from the physics and physiology of the standard causal picture, on which the argument depends. Although, as we have seen, the proper concerns of a *strictly-conceptual* thesis have been variously understood as words, meanings and lay usages, they could all be taken to disallow the evidence adduced by the argument from causation.

In summary, our argument from causation assumes the realism-claim, does not address the epistemological directness-claim, and is inadmissible with respect to any *strictly-conceptual* claim. Moreover it is a purely negative argument. Thus in the remainder of this chapter we focus on developing our argument from causation as an argument against the metaphysical directness-claim *factually* construed.

1.3.2 The structure of the argument

Smith notes that each of the four traditional arguments against direct realism attempts to find "some fact concerning the immediate object of awareness that does not hold of the normal [physical] object, or conversely". Thus, by the non-identity of discernibles, the immediate object of awareness is shown to be other than the physical object. (Smith, 2002, 8) On this basis the traditional arguments can be characterized in terms of the type of facts on which they draw: the arguments from illusion and hallucination on the facts concerning illusions and hallucinations, the argument from science on facts concerning the constitution of matter, and the argument from

causation on facts concerning the physical and physiological mechanism of perception. Since the "facts", on which the arguments depend, may be disputed, such facts are to be understood as putative facts which need to be substantiated.

For our purposes, we need to extend Smith's analysis and to think in terms of four argument schemata, into which particular instantiations of the arguments incorporate particular facts. (This will be important when we formalize the argument from causation in Section 1.3.3, and address the objections thereto in Chapter 2.) For example, particular instantiations of the argument from illusion schema might incorporate putative facts concerning double vision, partially immersed sticks or white walls in yellow light. Similarly there are different instantiations of the argument from causation. These various instantiations all draw their facts from the standard causal picture, that is a particular type of explanation of the process of perception in terms of a naturalistically understood causal chain of events. The content of the standard causal picture, and the particular facts which are drawn there from, are discussed in more detail below in Section 1.3.4 where we examine the premises of the argument.

An essential feature of our central thesis is that it does not maintain that metaphysical direct realism is false, only that it is incompatible with a particular naturalistic explanation of the perceptual process. Indeed we maintain that there are possible worlds in which the metaphysical directness-claim is true, but these are not worlds in which the process of perception can be explained in terms the standard causal picture. Indeed, that there are such possible worlds, follows from our understanding of the metaphysical directness-claim as a coherent *factual* claim.

In this respect we take "direct realism" more seriously than some of its proponents who appear to defend a claim which is true, but seemingly trivial.[34] We treat it as a *factual* thesis which makes claims which are true in some, but not all, possible worlds. On the other hand some of its

---

[34] Five decades ago, Armstrong talks of "some modern philosophers" who have taken "Direct Realism" to be true "but perfectly trivial". (Armstrong, 1961, p. xii) The same could be said today.

proponents, by claiming that the theory stands irrespective of the physics and physiology of perception in this our actual world, would seem to imply that the claims of metaphysical direct realism are true in all possible worlds. We will take up this issue when we address the "irrelevance objection" in Chapter 2 and the Gibsonian notion of direct perception in Chapter 3. For now it is sufficient to note our concerns as to the substance of any theory of perception which, by virtue of the arguments used in its defence, has the appearance of a necessary but seemingly trivial truth.

In accordance with this understanding of the issues, we do not present the argument from causation in the traditional form which takes the standard causal picture as a premise and draws the inference that direct realism is false. Rather we present the argument as an argument against the conjunction of direct realism and the standard causal picture. (Thus we use the term "conjunctive argument".) As to whether the perceptual process, in this our actual world, is such that it is explicable in terms of the standard causal picture, and as to the truth of the metaphysical directness-claim, we remain neutral. Thus our thesis would be uncontroversial and of limited interest, if it were not for the fact that much of contemporary analytic philosophy both espouses direct realism and accepts the standard causal picture. Contemporary analytic philosophy's allegiance to naturalism in general, and the standard causal picture in particular, is hardly controversial but is evidenced in the following paragraphs and Section 1.3.4.

"Naturalism" can be a slippery term. However the sense in which contemporary analytic philosophy espouses a naturalistic explanation of the perceptual process is that of "naturalism" in Dretske's "representational naturalism" (1997, p. xiii), Searle's "biological naturalism" (2004, p. 1), Wheeler's "Muggle constraint" (2005, p. 5), and in the particular context of perception, Martin's "experiential naturalism":

> *Experiential Naturalism*: our sense experiences, like other events or states in the natural world, are subject to the causal order, and in this case are thereby subject just to broadly

physical causes (ie including neuro-physiological causes and conditions) and psychological causes (if these are disjoint from physical causes.) (Martin, 2004, p. 39)

Here "naturalism" stands against a view which allows a different basis for the understanding of physical-physical causation and psycho-physical causation. On this understanding, very few contemporary analytic philosophers would explicitly reject naturalism by proposing explanations of the perceptual process which are discontinuous with natural science; contrast the Early Moderns (Stroud, 2008).[35] Foster is a rare exception who questions the assumption that "forms of causality which are operative in the physical realm will be operative in the psychophysical realm as well." (2000, p. 37)[36] Tellingly, Foster argues for idealism.

The notion of a naturalistically understood "causal chain of events" underpins the standard causal picture in terms of retinas, optic nerves and brains, which came to the fore in the seventeenth century, has been taken up by the scientifically informed layman, and is fleshed out in mainstream (as opposed to Gibsonian) cognitive science. That the perceptual process is, in principle, explicable in terms of the standard causal picture is now rarely questioned within the (analytic) philosophy of perception. Indeed objections to the argument from causation are usually preceded by an acknowledgement to this effect. For example:

> It's wise for Direct Realists to concede that for humans, and for percipients physiologically like us in the actual world, perception involves a long and complex causal series of events, and that perception is indeed dependent upon the condition of the eyes, of the optic nerve, and of the brain, upon the nature of the intervening medium, and so on. (le Morvan, 2004)

---

[35] In the Early Moderns God played an essential part in the perceptual process. "Descartes thought that human knowledge cannot be accounted for without a benevolent, omniscient, and omnipotent God who guarantees the truth of what human beings clearly and distinctly perceived to be true. For Berkeley, God's agency is the only active force there is in the world of things we perceive and know about. .... Even Lock relied on a benevolent agent as the ultimate source of those cognitive faculties that all that human beings need to get along in the world they find themselves in." (Stroud, 2008, p. 23)

[36] "The idea that earlier events [involving the perceived object] may have a direct influence on what happens after a temporal interval – an interval after which any record of the relevant features of those events has been lost – seems very strange and hardly credible." (Foster, 2000, p. 37)

The essential features of the standard causal picture, and the facts that can be drawn there from, are discussed further in Section 1.3.4 where we examine the premises of our formalized argument.

1.3.3 The logic of the argument

Here we formalize the argument from causation as an argument against the conjunction of the metaphysical directness-claim and the standard causal picture; it is in this form that the argument is needed to defend our central thesis. Although the structure of the argument is relatively simple, the formal presentation serves to make explicit its conjunctive form and the distinction between the general argument schema and particular instantiations thereof. It also serves to expose the premises for examination in Section 1.3.4.

Definition of terms:

D: Metaphysical directness-claim (The immediate object of awareness in standard perceptual situations is one and the same as, is identical to, a normal physical object.)

P: Weak Phenomenal Principle (In veridical perception, if there sensibly appears to a subject to be something which possesses a particular sensible quality then there is something of which the subject is aware which does possess that quality.)

N: Naturalistic standard causal picture (The perceptual process is such that it can be explained in terms of the standard causal picture.)

I: Indiscernibility of identicals (If X and Y are identical, they are indiscernible.)

X: Discernibility claim (The physical object and the immediate object of awareness are discernible.)

F_n: Some particular fact, drawn from the naturalistic standard causal picture, concerning the immediate object of awareness which does not hold of the physical object.

At the top level the argument can be formalized thus:

I	Premise 1
P	Premise 2
$P\&I\&N \rightarrow \neg D$	Premise 3

$\neg (D\&N)$

This is useful in making clear the overall structure of the argument as an argument against the conjunction of D&N, rather than against D given N. As we have noted, the argument is modest in this respect. It does not claim that metaphysical direct realism is not true, only that it is incompatible with an explanation of perception in terms of the standard causal picture.

However all the work is done by Premise 3 and further analysis is needed. We proceed in two stages. First we make explicit the inference from discernibility to the rejection of the metaphysical directness-claim. To this end we replace Premise 3 with Premises 4 and 5.

$(P\&N) \rightarrow X$	Premise 4
$(X \&I) \rightarrow \neg D$	Premise 5

Premise 4 asserts the implication between the standard causal picture and the discernibility of the immediate object of awareness and the normal physical object. Premise 5 asserts the implication between the discernibility of the immediate object of awareness and the normal physical object, and the falsity of the metaphysical directness-claim. The argument can now be presented as:

I   Premise 1

P   Premise 2

$(P\&N) \to X$ Premise 4

$(X \&I) \to \neg D$ Premise 5

$\neg (D\&N)$

This is a useful clarification, but Premise 4 needs to be unpacked. Why should it be supposed

that the standard causal picture implies the discernibility of the immediate object of awareness

and the normal physical object? Here we draw on our understanding of the argument from

causation as an argument schema into which particular instantiations of the argument incorporate

particular facts drawn from the standard causal picture. We need to distinguish between the

standard causal picture and the particular facts, drawn from that picture, on which particular

instantiations of the argument are based. To this end we use the notation F_n. The use of a

subscript shows that reference is being made to one of a number of particular facts. On this basis

we can replace Premise 4 with Premises 6 and 7.

$(P\&N) \to F_n$ Premise 6

$F_n \to X$ Premise 7

Premise 6 asserts the implication between the standard causal picture N and a particular fact F_n

concerning the immediate object of awareness which does not hold of the physical object. Note

that the claim that there is an immediate object of awareness in the veridical case, without which

the particular fact could not be asserted, is made explicit in Premise 2. Premise 7 asserts the

implication between the particular fact F_n and the discernibility of the immediate object of

awareness and the normal physical object.

Thus the argument from causation can be presented as:

I   Premise 1

P   Premise 2

$(P\&N) \rightarrow F_n$ Premise 6

$F_n \rightarrow X$ Premise 7

$(X\&I) \rightarrow \neg D$ Premise 5

$\neg (D\&N)$

The logic is clearly valid, as shown for completeness below. However the premises require closer examination.

I, P, (P & N) →Fn, Fn →X, (X & I) →¬ D : ¬ (D&N)

1	I	Premise 1	{1}
2	P	Premise 2	{2}
3	(P & N) → Fn	Premise 6	{3}
4	Fn → X	Premise 7	{4}
5	(X & I) → ¬ D	Premise 5	{5}
6	N	assumption for CP	{6}
7	P&N	&I 2,6	{2,6}
8	F_n	MP 3,7	{2,3,6}
9	X	MP 4,8	{2,3,4,6}
10	X&I	&I 1,9	{1,2,3,4,6,}
11	¬ D	MP 5,10	{1,2,3,4,5,6}
12	N→ ¬ D	CP 6,11 [discharges 6]	{1,2,3,4,5}
13	D & N	assumption for RAA	{13}
14	D	& E 13	{13}
15	¬ N	MT 12, 14	{1, 2,3,4,5,13}
16	N	& E 13	{13}
17	N & ¬N	&I 15, 16	{1, 2,3,4,5,13}
18	¬(D&N)	RAA 13,17	{1, 2,3,4,5,}

1.3.4 The premises of the argument

We take Premise 1, the indiscernibility of identicals, and thus the non-indentity of discernibles,

to be uncontroversial. As Smith puts it:

> I construe this law[37] here as the principle of "the indiscernibilty of identicals": if X and Y
> are one and the same, there can be no difference between them – since they are one.
> (2002, p. 275)

Similarly we take Premise 5, $(X \& I) \rightarrow \neg D$, to be uncontroversial. It follows from the definition

of the metaphysical directness-claim as an identity claim; if the immediate object of awareness

and the normal physical object are discernible, then, given 1, they are not one and the same, thus

the identity claim is false.

As discussed in Section 1.2, since Premise 2, the Weak Phenomenal Principle, makes no claims

about an immediate object of awareness in the illusory or hallucinatory cases, it is robust and

rarely disputed explicitly. However, it could be rejected by a thoroughgoing reductive

representationalist and cannot be taken to be uncontroversial. It will require further examination

when we address the objections to the argument in Chapter 2.

Here we focus on Premises 6 and 7 which are peculiar to the argument from causation. First, we

examine the naturalistic standard causal picture N which underpins the argument from causation

schema and, second, certain of the facts, F_n, which follow from N, and form the basis of

particular instantiations of the argument. In what follows we will make use of Vision's locution

"visual episode" (1997, p. 22) in order to refer to the perceptual experience without the need to

---

[37] Here Smith refers to "Leibniz's Law", though this is more usually taken to be the more controversial claim that
indiscernibles are identical. (Forrest, 2012) However Robinson aligns with Smith and in the same context refers to
Leibniz's Law as saying that "*a* and *b* cannot be identical if one possesses a property …that the other lacks." (1994,
p. 81)

call on Premise 2, the Weak Phenomenal Principle, to sanction the notion of an "immediate object of awareness".

We need to make explicit certain features of the standard causal picture, N, which explains perception in terms of a naturalistically understood causal chain of events involving light, retinas, optic nerves etc. Although they are unlikely to be controversial within contemporary analytic philosophy, the implications thereof are not always recognized in the philosophical literature and the argument from causation is thereby weakened. Moreover, when these features of the standard causal picture are made explicit, the possibility of non-standard causal pictures which are more, or less, conducive to the notion of direct perception is brought to the fore, and thus direct realism is more clearly understood as a *factual* thesis which is true in some but not all possible worlds. We will consider three features of the standard causal picture, with respect to which it stands opposed to alternative non-standard causal pictures which are arguably more sympathetic to the notion of direct perception.

Firstly the standard causal picture, in terms of light reflected from objects into the eye and signals travelling down the optic nerve to the brain etc, constitutes an intromission theory of visual perception, rather than an extramission theory. In contrast to the intromission theory in which something impinges on the eye, in the extramission theory visual rays emanate from the eye. For example, according to Euclid, a leading proponent of the extramission theory, "Rectilinear rays proceeding from the eye diverge infinitely [and] those things are seen upon which the visual rays fall…" (Gross, 1999, p. 58) Theories of both types, and combinations of the two, are to be found in the Greeks and the debate continued until the Early Modern Period when the extramission theory was finally rejected as being inconsistent with the physics and

physiology of the perceptual process.[38] Interestingly, studies have shown that extramission theories of perception are widely held by children and by some college students. (Gross, 1999) (Winner, et al., 2002) Sheldrake, controversially, promotes an extramission theory in his account of "the sense of being stared at". (2005) More significantly from our perspective, as discussed in Chapter 3, Gibson suggests the senses are analogous to "tentacles or feelers" rather than, as in the standard causal picture, the "initiators of signals" (1968, p. 5), and advertises his avowedly radical theory of direct perception "as a return to the optics of Euclid and Ptolemy." (Gibson, 1974)

Secondly, the standard causal picture stands in opposition to an understanding of the perceptual process which, although it is in accordance with our pre-critical intuitions and conducive to theories of direct perception, became highly implausible in the light of seventeenth century science. Stout's uses the particularly descriptive term "penny-in-the-slot theory" when he takes issue with Alexander:

> According to him [Alexander], when we perceive objects as external to our sense organs the sensum is itself a feature of the external object having an existence and nature of its own, quite independent of its entrance into our sense experience, whereby it becomes a sensum. The physical and physiological conditions of perception simply serve to unveil it. They put a penny in the slot and so, as it were, remove a screen which would otherwise have hidden it from the percipient. (Stout, 1922, p. 387)

Stout's reference to penny in the slot machines is now dated, but the modern day example of the coin operated telescope would do equally well. In this non-standard causal picture, the perceptual process is analogous to the mechanism by which the coin is registered and the telescope's aperture is unblocked, in contrast, say, to the mechanism of a closed circuit TV surveillance system in which transduction and signal processing are involved. The key point is that, in

---

[38] Feynman puts it well in discussing the physics of vision. "There have been many theories, but it finally settled down to one, which is that there is something which enters the eye – which bounces off objects into the eye. …. it is almost impossible for us to realize that very intelligent men have proposed contrary theories – that something comes out of the eye and feels for the object, for example." (1963, 26 -2)

distinction to the standard causal picture, the perceptual process plays a merely passive role and serves only to enable, or switch on, our perceptual faculty. More recently Maund quoting Burnyeat refers to a "window model of perception" which has been attributed to a number of "historical philosophers" according to which "like opening a window and looking out, the act of perceiving reveals the thing, perceived "as it really is"." (Maund, 2003, p. 8)

The key point is that though they may now be implausible, there are non-standard causal pictures in which the process of perception serves only to unveil, or reveal, the external object. And here again, we will find suggestions to this effect in Gibson's theory of direct perception in terms of "information pickup" by "psychological resonance". (1968, p. 271)

Thirdly, the standard causal picture assumes local rather than non-local causation. The notion of local causation aligns closely with Hume's first rule "by which to judge of causes and effects" that is "The cause and effect must be contiguous in space and time". (Schliesser, 2008). Conversely, Lockwood defines non-local causation, that is "action at a distance", as "action without any dynamical contact or spatio-temporally intervening process." (1989, p. 203) If non-local causation were to be operative in the mechanism of perception, say in terms of action at a distance between objects and brains, explanations in terms of the intermediary processes would be circumvented[39]. However, the claim that non-local causation is operative at the level of the perceptual processes is rejected within mainstream cognitive science and finds little support within contemporary analytic philosophy. Martin puts it well:

> It may well not be excluded entirely *a priori* that there should be action at a distance within the neurological realm, but few neuroscientists with the hope of serious funding would pursue the hypothesis seriously.[40] (2004, p. 55)

---

[39] The notion of a causal chain of events, within which causation is local, underpins contemporary explanations of the perceptual process in terms in terms of say, the successive contribution of visual areas V1, V2, V3 etc The same can be said for explanations at other levels, say, in terms of networks of, axons, synapses and dendrites. (Bruce, et al., 2003)

[40] Martin may well be correct with respect to the funding of neuroscience, but the situation is very different in physics where considerable funding has been given to high profile experiments to test the predictions of Bell's Theorem (d'Espagnat, 1979), which is highly significant to the debate about non-local causality. However it is

Although the notion of action at a distance would support Martin's defence of direct realism[41], he rejects this easy way out as incompatible with his commitment to naturalism.[42] Similarly, as Edelman puts it with respect to formulating his theory of consciousness, "I allow no spooks … no action at a distance." (1992, p. 113)

However here again, although there is no support for the notion of non-local causation within mainstream cognitive science, some of Gibson's intellectual successors are sympathetic. For example, Turvey questions the assumption of local causation when he suggests that "Advances in the theory of direct perception should follow from an explicit grounding of the theory in the physical assumptions of the 21st century rather than the 17th century." (2003) Moreover, though controversial, the suggestion that non-local causation, and perhaps quantum mechanical effects as yet unknown to physics, might play an essential role in cognition has been more widely discussed, notably by Penrose (1989) (1994). Indeed, drawing on references to quantum mechanics, Hodgson (1996, p. 18) takes issue with Edelman maintaining that his "exclusion of action at a distance is unsupported by any argument."

It must be emphasized that we are not arguing for these non-standard causal pictures. Indeed we accept that there may be good reasons, here in our actual world, for endorsing the standard causal picture. The key point is to contrast the standard causal picture with alternative causal pictures which are more sympathetic to the notion of direct perception, and with regard to which the argument from causation could not be framed. Moreover, as intimated above, and discussed

---

certainly true that mainstream cognitive science rejects the claim that non-local causation is operative at the level of the perceptual process.

[41] Here Martin is defending his disjunctivist objection to the argument from hallucination against the "doctrine of 'Same Cause Same effect'" according to which "where any situations involve the same proximate causal conditions then the very same kind of effects will occur." (2004, p. 55)

[42] Whether or not an explanation of perception in terms of action at a distance would count as a "naturalistic" explanation is moot. It is debatable how far an explanatory framework can depart from that of contemporary natural science whilst maintaining continuity. We will take up this issue in Appendix A where we address such radical departures from the mainstream cognitive science.

more fully in Chapter 3 and Appendix A, there are suggestions of all three non-standard causal pictures in Gibson's avowedly radical theory of direct perception and the ideas of his intellectual successors.

Having made explicit the key features of the standard causal picture, N, we can now formulate particular instantiations of the argument from causation schema by drawing on N to assert particular facts, F_n. It is at Premise 6, $(P\&N) \rightarrow F_n$, where particular instantiations of the argument from causation draw on the standard causal picture N, along with the Weak Phenomenal Principle, P, to assert a particular fact, F_n, concerning the immediate object of awareness that does not hold of the normal physical object, or vice versa. Here it is useful to return to Smith's brief survey across all four traditional arguments against direct realism and the facts on which they draw:

> Concerning the immediate object of awareness and the normal object: (a) one possesses a genuine attribute that the other lacks; (b) one bears a genuine relation to another item which the other does not; (c) one exists at a place that the other does not; (d) one exists at a time when the other does not; (e) one exists and the other does not. (Smith, 2002, p. 9)

Smith is only concerned with the arguments from illusion and hallucination which he identifies as corresponding to (a) and (e) respectively. However, cases (b), (c) and (d) can be seen as providing the basis for the various instantiations of the argument from causation: for example, an argument based on the type-b fact that the character of the immediate object of awareness is, whilst that of the normal object is not, dependent on functioning of our brains and sensory systems; an argument based on the type-c fact that the normal object and immediate object are located at different points on a causal chain; and an argument based on the type-d fact that the immediate object of awareness may exist at a time when a distant normal object does not.

Developing the first example, of the type-b fact, Premise 6 of the argument schema, $(P\&N) \rightarrow F_n$, can be instantiated thus as Premise 6_1:

(P&N) → (The character of the immediate object of awareness is, whilst that of the physical object is not, dependent on the functioning of our brains and sensory systems)

With respect to this premise, two issues need to be clarified. Firstly, the "character" of the immediate object of awareness is to be understood as "phenomenal character." As discussed in Section 1.2, we take "phenomenal character" to be a property of the "immediate object of awareness": the immediate object of awareness is the bearer of phenomenal character. [43] Thus a property of the immediate object of awareness is dependent on the functioning of our brains and sensory systems, whilst no property of the physical object is so dependent.

Secondly we need to clarify the notion of dependency. Here "dependent" is to be taken to mean that some changes in the functioning of our brains and sensory systems result in changes in the character of the immediate object of awareness. This is a relatively modest claim. It does not rely on the principle of same proximate cause same immediate effect which might be taken to be unsupported by the standard causal picture[44], nor is it necessarily opposed to an externalist understanding[45] of mental content. However the claim does rely on the standard causal picture's opposition to a "penny in the slot" theory of perception. If the our brains and sensory systems served only to unveil the normal physical object, as in Stout's "penny in the slot" analogy discussed above, this particular instantiation of Premise 6 could not be maintained.

---

[43] The time-gap argument and Robinson's version of the argument from causation, as discussed in Section 2.2, draw, respectively, on type-d and type-c facts in Smith's taxonomy. These instantiations do not depend on this particular understanding of "phenomenal character".

[44] For example Fish (2009, p. 116), in defending the disjunctivist response to the argument from hallucination against the principle of same proximate cause same immediate effect, argues that, although it is often assumed that phenomenal character supervenes on the state of our brains ("total local supervenience"), the assumption is unfounded and not supported by the empirical evidence. However, although the argument from hallucination (in some forms) does draw on this assumption, the above argument from causation does not.

[45] "*Externalism* with regard to mental content maintains that in order to have certain types of intentional mental states (e.g. beliefs), it is necessary to be related to the environment in the right way." (Lau & Deutsch, 2013)

Following through this example, Premise 7 of the argument schema, $F_n \rightarrow X$, is instantiated thus as Premise 7_1:

> (The character of the immediate object of awareness is, whilst that of the physical object is not, dependent on the functioning of our brains and sensory systems) $\rightarrow X$

We take Premise 7_1 to be uncontroversial. If changes in the functioning of our brains and sensory systems result in changes in the character of the immediate object of awareness but not in the character of the physical object, then the immediate object of awareness and the physical object are discernible.

So here we have a particular instantiation of the argument from causation schema:

I	Premise 1
P	Premise 2
$(P\&N) \rightarrow F_1$	Premise 6_1
$F_1 \rightarrow X$	Premise 7_1
$(X \& I) \rightarrow \neg D$	Premise 5

$\neg (D\&N)$	

Loosely put, this particular instantiation of the argument schema can be expressed as follows. Since, according to the standard causal picture, the character of the immediate object of awareness is, whilst that of the physical object is not, dependent on the functioning of our brains and sensory systems, it follows that, contrary to metaphysical direct realism's directness-claim,

the immediate object of awareness cannot be one and the same as the physical object. Thus the standard causal picture and metaphysical direct realism cannot both be true.

1.4 Conclusions

1. As a preliminary to examining the argument from causation we need to disambiguate the thesis which is at issue, and particularly to distinguish between realism v directness claims, metaphysical v epistemological claims and *factual* v *strictly-conceptual* claims, and to separate negative arguments against direct realism from positive arguments for some alternative theory of perception.

2. Drawing on these distinctions, we present the argument from causation as an argument against the conjunction of the standard causal picture and the metaphysical directness-claim *factually* construed. It is in this form that the argument is required to defend our central thesis. (We do not present the argument as an argument against epistemological direct realism or any *strictly-conceptual* thesis, nor as a positive argument for some alternative theory of perception.)

3. We maintain that, presented as above, the argument from causation is valid and the premises are robust; though not entirely uncontroversial, they would be widely accepted within much of contemporary analytic philosophy. Moreover, since only the veridical case is at issue, the argument avoids the disjunctivist objection altogether and seriously weakens the representationalist objection. Thus it presents a particularly serious challenge to metaphysical direct realism which would need to be answered.

4.   Nevertheless, as we will see, the argument from causation has been dismissed lightly as irrelevant or confused in much of the philosophical literature. These objections are addressed in Chapter 2.

Chapter 2: Defusing the standard objections

This chapter challenges the standard objections on the basis of which the argument from causation has been dismissed lightly as irrelevant or confused in much of the philosophical literature. As we have noted, and further exemplified below, the literature is often compromised by ambiguities in both the argument under examination and the thesis of "direct realism" which is being defended. Thus objections which are valid in the context of a weakly formulated or loosely targeted argument are taken to be conclusive. Here we defend the argument formalized in Chapter 1, that is an argument targeted on the conjunction of the standard causal picture and the metaphysical directness-claim *factually* construed.

We start, in Section 2.1, with Armstrong's critique of the argument in his seminal defence of direct realism. (1961) Here we find three types of objection: type-1, that the argument is self-refuting since the conclusions undermine the premises; type-2, Armstrong's "fundamental criticism", that the considerations adduced by the argument, with respect to the physical and physiological mechanisms of perception, are irrelevant to the standing of direct realism; and type-3, that the argument depends on the controversial Phenomenal Principle. We then consider Armstrong's treatment of the time-gap version of the argument, which casts further light on the nature of the thesis he is defending.

Section 2.2 addresses the more recent literature. Here we find examples of the same three types of objection, but again, as in Armstrong, the argument is most often dismissed lightly on the basis of type-2 irrelevance objections. We maintain that although these objections might serve to defend the epistemological directness-claim, or some *strictly-conceptual* claim, they have no purchase against an argument which is clearly targeted on the metaphysical directness-claim *factually* construed. Moreover, type-1 objections, that the arguments is self refuting, cannot be applied against our conjunctive version of the argument which claims only that the standard

causal picture and the metaphysical directness-claim cannot both be true. Only the type-3 objections to our call on the Weak Phenomenal Principle remain as a threat to the argument formalized in Section 1.3. Section 2.3 summarizes our conclusions.

2.1 Armstrong's critique

Published around the time of the resurgence of interest in the philosophy of perception, Armstrong's monograph *Perception and the Physical World* (1961) is the first systematic attempt to defend direct realism against all four traditional arguments. (Unless otherwise stated, references in Section 2.1 are to this monograph.) His treatment of the arguments is still of considerable interest since it introduces the themes which have been taken up in the later discussions. Moreover his somewhat dismissive response to the argument from causation is typical of much of the later work, and his claim that the argument "is to be rejected out of hand" (p. 144) provides a good point from which to examine the more recent literature.

We have noted that much of the literature is compromised by ambiguities in the thesis which is being defended, particularly with respect to a failure to distinguish between metaphysical and epistemological claims and between *factual* and *strictly-conceptual* claims. Armstrong's defence of "direct realism" against the argument from causation is a case in point. However, it seems clear that Armstrong intends to defend, at least[46], the metaphysical directness-claim. He introduces the theories of perception as answers to the seemingly metaphysical question "What is the *direct or immediate* object of awareness when we perceive?", to which direct realism answers "that the immediate object of awareness is never anything but a physical existent …" (p. xi) Here Armstrong uses a similar form of words to that used by Smith which we have taken as our definition of the metaphysical directness-claim, that is "the immediate object of awareness in standard perceptual situations is a normal physical object." (Smith, 2002, p. 7) Thus, in this

---

[46] Armstrong could also be construed as defending epistemological direct realism. See particularly Chapter 8 "The Argument from Verification" (Armstrong, 1961, p. 101)

respect at least, we can take Armstrong to be defending the thesis of metaphysical direct realism against which we are targeting the argument from causation.

On the other hand, the ambiguities with respect to the conceptual and factual content of his thesis are more difficult to resolve. Armstrong's introductory remarks suggest that he is proposing to defend "direct realism" as a type of *strictly-conceptual* thesis which places a restriction on the admissible data:

> It is true that this [What is the direct or immediate object of awareness when we perceive?] is not a scientific question, to be answered by observation and experiment, but is a conceptual question to be answered by the means appropriate to the discussion of a conceptual question, viz philosophical argument. But it is just as real a question as, say, the conceptual question 'What is an unconscious wish?' (p. xii)

However, it is not at all clear what Armstrong means by a "conceptual question". Indeed it is unclear whether he means to say that "observation and experiment" is irrelevant or insufficient. If the former, he is defending a *strictly-conceptual* thesis, if the latter a *factual* thesis. As an exemplar of the type of question at issue, "What is an unconscious wish?" is unhelpful since it could be variously understood.[47] However, on the most natural interpretation, it appears to be intended as an exemplar of a question with respect to which "observation and experiment" are irrelevant rather than merely insufficient. On this basis Armstrong could be taken as defending a *strictly-conceptual* thesis and, as we will see, such an interpretation would shed light on his otherwise enigmatic version of the type-2 "irrelevance objection".

However, this deflationary understanding of the question at issue does not appear to do justice to the thesis which Armstrong is defending. Firstly, Armstrong distances himself from "some modern philosophers", presumably the Linguistic Philosophers of the 1950s, who interpreted his opening question in such a way that "Direct Realism has been taken to be true but perfectly

---

[47] For example, as a question about the neurological correlates of respectively conscious and unconscious wishes, a question framed in the context of Freudian Psychology, a question about the use of the concept of an "unconscious wish" as it is used by philosophically and scientifically unsophisticated English speakers, or a question about the meaning of and relationship between the terms "unconscious" and "wish".

trivial." (p. xii) Secondly, he takes the scientific data seriously, as is evidenced by his detailed discussion of the argument from science which he regards as a serious threat to his position. (pp. 155 - 171) Here Armstrong does not place any limitations on the admissibility of data. Lastly, he appears to go far beyond maintaining a *strictly-conceptual* thesis about, say, how perceptual concepts are employed by their lay users, and defends a *factual* thesis about the nature of this world in which we find ourselves. For example, he sums up the direct realist "view of perception and the physical world" which, he believes, has survived the arguments from illusion and hallucination as follows:

> The world is much as it looks to be, feels to be, smells to be, tastes to be, or sounds to be. At the same time the world exists quite independently of actual or possible looking, feeling, smelling, tasting or hearing. (p. 139)

This is a claim about how things are, here in our actual world, which may or may not be true. Moreover it is the claim with respect to which the Early Moderns take issue, say Reid, Locke and Berkeley, and it goes beyond a *strictly-conceptual* thesis of the type discussed in Section 1.2.3.

 Thus there remains a major ambiguity in the nature of the thesis which Armstrong is defending. His introductory remarks, with respect to the nature of the "conceptual question" he is addressing, suggest that he is defending direct realism construed as a *strictly-conceptual* thesis. However, the broader thrust of the monograph is at odds with any such deflationary construal and, on this basis, we will assume, provisionally at least, that he is defending the *factual* thesis which we are targeting. (We will need to revise this assumption in Section 2.2.)

Before focusing on Armstrong's critique of the argument from causation, we need to put this section of his monograph into the context of the whole, and particularly his treatment of the arguments from illusion and hallucination. He opens with a discussion of those considerations that have been taken to undermine the direct realist position and lead to indirect realism or phenomenalism. After devoting the first 80 pages to refuting both alternatives, he sets out to

defend direct realism as the one remaining option. Armstrong's counter to the arguments from

illusion and hallucination is, in our terminology, to reject the Phenomenal Principle, and block

the argument at the base case, on the basis of his reductive analysis of perception in terms of

belief acquisition. On this understanding:

> …sensory illusion is nothing but false belief, or inclination to false belief, that we are
> perceiving some physical object or state of affairs. (p. 87)

Clearly, if Armstrong's analysis of non veridical perception is accepted, the argument from

illusion/hallucination[48] is blocked at the base case.[49] If sensory illusion is nothing but an

inclination to false belief, the Phenomenal Principle must be rejected and there is no immediate

object of awareness in the sense that the argument requires. Having addressed the illusory and

hallucinatory cases, Armstrong goes on to presents the more radical thesis that even in the case

of "ordinary, veridical, perception":

> …perception is nothing but the acquiring of knowledge of, or, on occasions, the
> acquiring of an inclination to believe in, particular facts about the physical world, by
> means of our senses (p. 105)

He recognizes at the outset that this is a highly controversial thesis. The problem, of course, is

that veridical perception seems to be much more than acquiring beliefs about the physical world

by means of the senses. Armstrong anticipates this objection and accepts that veridical

perception involves, additionally, "a flood of up-to-date information about our environment",

"characteristic sensations in the organs of perception" and "knowledge of the means by which

this knowledge is acquired". (p. 112) However, he recognizes that the objectors will not be

satisfied with these minor concessions and that they will continue to argue that there is still more

---

[48] Armstrong does not always maintain this distinction and discusses both argument under the heading "The Argument from Illusion". (p. 19)

[49] Armstrong must block the argument at the base case, since he accepts the Common Factor Principle and cannot deploy the disjunctivist's objection to the spreading step. He talks of "the parallel we must maintain between sensory illusion and veridical perception." (1961, p. 83)

to perception, specifically the visual experience on which the beliefs or inclinations to believe are grounded.[50]

So here, in our terminology, Armstrong is rejecting the Weak Phenomenal Principle, and with it, the seemingly robust claim that there is an immediate object of awareness in the veridical case. These will be important considerations when we come to address Armstrong's objections to the argument from causation.

Having addressed the arguments from illusion and hallucination, Armstrong takes stock of the situation:

> But the common-sense view of perception and of the physical world has survived these assaults only to meet attacks from a different quarter, attacks which are as hostile to Phenomenalism as they are to Direct Realism, and which tend to reinstate some form of Representative theory [Lockean indirect realism]. These attacks draw their inspiration from the facts and theories of physics and physiology… (p. 140)

He goes on to introduce the "Argument from Causation" and the "Argument from Science":

> There seem to be two main lines of attack here, one based on the physical and physiological mechanisms of perception [the Argument from Causation], the other on theories of the constitution of matter [the Argument from Science]. (p. 140)

Armstrong's treatment of these two arguments is in sharp contrast. Although he dismisses the former as thoroughly confused, he believes that the latter "leads to most serious doubts about the common sense picture of the world that perception furnishes us with." (p. 140).[51] Indeed, in contrast to the arguments from illusion, hallucination and science which are all considered at length, "The Argument from Causation" is "rejected out of hand" after a mere three pages.

---

[50] The crux of his response is that "we are forced to reject" the more plausible alternative, that our perceptually acquired beliefs are "grounded on further 'perceptual experience'", because to do otherwise would lead to "unacceptable consequences", that is the acceptance of the indirect realist or phenomenalist theory of perception, both of which he has already rejected. (p. 117) Here, it seems, we are enjoined to accept direct realism as the "least worst" theory of perception even if it is at odds with the phenomenology of our perceptions.

[51] He gives considerable space to the argument from science, pages 155 -171, and concludes "But all this [his response to the argument from science]… is very tentative. In our discussion whether it is possible to reconcile a Direct Realist account of perception with a Realistic account of modern physics we have found difficulties that we have not been able to solve." (Armstrong, 1961, 190)

However this avowedly decisive rejection needs to be qualified, since Armstrong then devotes a further nine pages to "The Argument from the Time-Gap" which, on our analysis, is a particular instantiation of the argument from causation schema. We will follow the order of Armstrong's presentation and consider first his objections to "The Argument from Causation" and then, under a separate heading, his more qualified objections to "The Time-Gap Argument". His treatment of the latter will help us to better understand the nature of the thesis which he takes to be at issue in the former.

With respect to his rejection of "The Argument from Causation" as "thorough-going confusion", two questions arise. Firstly, on what basis does he reject the argument? Secondly, why does he consider its failings to be so obvious? The second question will have a significant bearing on the first.

Armstrong introduces the argument from causation thus:

> The argument bases itself on the scientific discovery that, before we can perceive anything, a chain of processes must begin[52] in the object, travel through our sense organs, and reach the brain. … Until these complex processes occur, perception cannot occur. It is inferred from this that there can be no *immediate* perception of the physical object or happening that we say is perceived. The only possible *immediate* object of perception is the last link in the chain of processes. This link is usually identified with a sense impression, whose immediate cause is a happening in the brain. (p. 141)

Here we have an allusion to, rather than a statement of, the argument from causation. Significantly, it makes no reference to the basis of the key inference from "complex processes" to "no *immediate* perception". However it does serve to identify the traditional argument which draws on the physical and physiological mechanism of perception. Moreover, since what is at issue is the identity of the "physical object" and the "*immediate* object of perception", it seems clear that Armstrong takes the argument to be targeted on metaphysical direct realism's directness-claim. Furthermore, although the ambiguity remains, we will assume, provisionally,

---

[52] Nothing rests on Armstrong's suggestion that the "chain of processes" begins in the "object". As discussed in the context of "The Causal Theory of Perception" (Vision, 1997), the light source might be taken to be part of the relevant causal chain.

that in rejecting the argument he is defending metaphysical direct realism as a *factual* rather than merely a *strictly-conceptual* thesis.

For the purpose of analysis and comparison with the more recent literature it will be useful to group Armstrong's objections into three types.

2.1.1 Type-1 objections: the argument is self-refuting

Armstrong's first criticism is that the argument leads to an obvious falsehood, that the "immediate objects of perception" are "brain states". He concludes:

> The Argument from Causation will then have to make brain-states the immediate objects of perception, which is a *reductio ad absurdum* of the argument. The only alternative would be to find some third candidate for the "immediate object of perception", but there seems to be no such candidate in sight. (p. 142)

This objection is questionable on a number of counts.[53] However, even if it was effective against a version of the argument from causation which makes positive claims about the immediate objects of awareness, it has no purchase against our more modest argument which makes only the negative claim that the immediate object of awareness cannot be one and the same as the normal physical object.

However the charge that the argument is self-defeating is less easily avoided in Armstrong's next criticism, which claims that the argument's conclusions undermine the premises on which it is based:

> Critics of the Argument from Causation have regularly pointed out that, if the conclusion of the argument is true, viz. that no physical state of affairs (with the possible exception of brain-states) are ever immediately perceived, then we have no immediate evidence for the facts which are adduced as premises of the argument, that is, the behaviour of light and sound waves, of sense organs and of our nervous system. (p. 142)

---

[53]Firstly, Armstrong's claim that there are no "other candidate in sight" rests on his thoroughgoing reductive analysis of perception and hence the rejection of "sense impressions" as suitable candidates. Secondly the suggestion that the immediate objects of perception are "brain states" is too close to the implications of theories that have actually been proposed to allow it (the assertion) to be simply rejected as an obvious falsehood. For example, Russell's notorious assertion, "I should say that what the physiologist sees when he looks at a brain is part of his own brain, not part of the brain he is examining." (1927, p. 383)

He notes that this objection is a version of the "familiar argument" that he brought against the indirect realist theory of perception earlier in the monograph:

> … if the theory were true, we could have no good reason for postulating the existence of physical objects to be the causes of our immediate perceptions. Here it [the objection] comes with the extra sting in its tail that, that if we accepted the Representative theory [Lockean indirect realism] on the basis of the Argument From Causation, we should have to become skeptical about the very evidence which was adduced to prove the Representative theory. (p. 142)

Here we agree with Armstrong. Indeed, as discussed in Section 2.2.1, we are more sympathetic to this type of objection to the traditional form of the argument than are some of the more recent commentators. However, it has no purchase against our conjunctive version. That Armstrong does not pursue this type-1 objection is perhaps explained by his reliance on the following.

2.1.2 Type-2 objections: the causal considerations are irrelevant

We now turn to Armstrong's "fundamental criticism" in comparison to which the above is "but a preliminary debating point". The causal considerations which the argument brings to the table are irrelevant to the standing of the thesis of direct realism that he is defending:

> The argument has confounded two quite distinct things: (a) perceiving an X or perceiving that an X is Y; and (b) the causal conditions which bring about this perception. It cannot be denied that perception occurs when, and only when, a certain very complex process begins in the object perceived and ends in the brain. But what warrant have we for identifying this with perceiving? May it not simply be the necessary, or even the necessary and sufficient, *preconditions* of perception: that which must occur if perception is to occur, but which is not to be identified with perception itself? The argument has done nothing to show that there is any identity here. And the fact that we know very well what perceiving is long before we know anything about these complex processes, suggests very strongly indeed that the two are *not* to be identified, however closely they may be related. (p. 142)

If, as Armstrong maintains, the argument from causation did confound (a) and (b) this would be a serious objection. Indeed it would explain why Armstrong is so confident that the argument is not only invalid, but thoroughly confused and "to be rejected out of hand." However the argument does not assume or assert that there is an "identity" between (a) and (b). Moreover, it is very difficult to understand why Armstrong should suppose that the argument makes what is an

obviously fallacious identity claim. How could any act, an act of "perceiving" in this case, be taken to be one and the same as "the causal conditions which bring about" that act?

Nevertheless, as Armstrong sees it, the argument's reliance on this supposed identity is clearly the issue at the core of his "fundamental criticism". Indeed the reference to identity is continued in Armstrong's somewhat puzzling conclusion to the section: "The beating of the light-waves on my eyes and brain makes me see, but seeing is not **identical** [emphasis added] with the beating of the light-waves." (p. 143) The implication is clearly that the argument from causation wrongly assumes or asserts such an identity. But the charge appears to be without foundation.

The force of Armstrong's "fundamental criticism" is still more difficult to understand when it is directed at our formalized argument, and the focus is on particular instantiations of the argument schema which draw on the standard causal picture to assert particular facts concerning the immediate object of awareness which do not hold of the normal physical object. Here, there is clearly no suggestion of an identity between (a) and (b), but rather that facts about (b), "the causal conditions", have implications for the nature of the "perceiving" involved in (a), most importantly whether it constitutes direct or indirect perceiving in the sense of the metaphysical directness-claim. Thus, for the objection to go through Armstrong would need to claim, not that (b) is distinct from (a), which it clearly is, but to claim that (b) has no bearing on (a). Indeed Armstrong could be read as attempting to justify the latter claim.

Armstrong's suggestion that "a certain very complex process" which "begins in the object perceived and ends in the brain" serves only as the "*preconditions* of perception" could be taken as supporting a non-standard causal picture, that is the "penny-in-the-slot theory" as discussed in Section 1.3.4, within which the perceptual process merely serves to unveil the physical object,

and the character of the corresponding visual episode is independent of the physics and physiology of the perceptual process. However, although this interpretation might serve to explain the basis of Armstrong's "fundamental criticism", it seems most unlikely that he would endorse any such non-standard causal picture. Indeed his commitment to the standard causal picture is evidenced by his efforts to reconcile direct realism with "the physicist's account of reality [as] a literal picture[54] of the physical world" in his lengthy discussion of the argument from science. (pp. 165 - 191) Thus we see no reason to suppose that Armstrong is, as in the case of Gibsonian cognitive science discussed in Chapter 3, defending the notion of direct perception at the expense of the standard causal picture.[55]

There appears to be just one further interpretation. Perhaps the real force of Armstrong's "fundamental criticism" is to be found in his above reference to "perception itself" and "the fact that we know very well what perceiving is long before we know anything about these complex processes". Here there are similarities with Price's claim that "we want to know what seeing and touching themselves *are"* and, explicit in Price, that this "question lies outside the sphere of Science altogether." (1950, p. 2) Moreover, if direct realism is construed as a thesis about "what perception is", which stands independently of our knowledge of the "causal conditions which bring about perception", that is stands independently of considerations with respect the physical and physiological mechanism of perception in this our actual world, then the irrelevance objection goes through. Indeed the argument from causation could, as Armstrong proposes, be

---

[54] He discusses but ultimately rejects scientific phenomenalism, although he recognizes it might provide a defence of direct realism against the Argument from Science. "So perhaps the Direct Realist can evade the conclusions of the Argument from Science by arguing these entities of which the physicist speaks, molecules, atoms, electrons, protons, light-waves, etc, are simply theoretical concepts." (Armstrong, 1961, p. 157)

[55] However, he might be taken as questioning the standard causal picture in the following passage. "Now it is very easy to think of the physical process from object to brain as a chain of which we can only grasp the last link in our brain or mind. We conclude that the last link is the only immediate object of perception. But of course the model only has to be exhibited clearly to be rejected." (Armstrong, 1961, p. 144) Unfortunately, as presented, this model might be rejected for a number of reasons and, since Armstrong does not state his particular objection, it is difficult to draw conclusions as to exactly what he is rejecting.

"rejected out of hand". On this understanding Armstrong's direct realism would be a *strictly-conceptual* thesis which, like the "The Causal Theory of Perception" discussed in Section 1.2.3, limits the data, that may count for or against, to that available to lay users of the concept of perception. However, as we have noted, this deflationary construal of the thesis does not do justice to the broader thrust of Armstrong's monograph.

Thus Armstrong's objection that the argument confuses "perceiving an X" and "the causal conditions which bring about this perception" remains enigmatic. It might have been understood on the basis of a rejection of the standard causal picture, or a deflationary construal of the thesis at issue, but neither suggestion is supported by the main thrust of the monograph. We will return to Armstrong's "fundamental criticism" when we address his treatment of the time-gap argument in Section 2.1.4, and more recent versions of type-2 objections in Section 2.2.2. (There we will conclude that, although considerable ambiguity remains, it seems that Armstrong is best understood as defending a form of *strictly-conceptual* thesis.)

2.1.3 Type-3 objections: the argument depends on the Phenomenal Principle

After disposing of the argument from causation on the basis of his "fundamental criticism", Armstrong sets out to explain why, given what he takes to be its obvious shortcomings, the argument from causation "ever had any persuasive power in the first place." (p. 143) Here he brings to the fore his claim that the argument from causation is dependent on, and subordinate to, the argument from illusion. As he puts it, "the Argument [from causation] makes its living by taking in the dirty washing of the Argument from Illusion." (p. 143) Thus after restating his objection to the argument from illusion he concludes:

> So the Argument from Causation only presents us with problems if the problems of the Argument from Illusion are unsolved. Once the latter *are* solved, the former present no independent difficulties. (p. 143)

In Section 1.1 we argued that, to the contrary, the argument from causation is not dependent on the argument from illusion and is, in important respects, the more robust of the two. In the terminology of our analysis, whereas Armstrong maintains that the argument from causation depends on the argument from illusion to establish the Phenomenal Principle, we maintain that both arguments assume the Phenomenal Principle, though the argument from causation assumes only the less controversial Weak Phenomenal Principle which makes no claims about an immediate object of awareness in the problematic illusory or hallucinatory cases.

Nevertheless, it could be maintained that, independent of any reference to the argument from illusion, Armstrong's reductive analysis of perception as belief acquisition, which he defends at length earlier in the monograph (pp. 105 -122), constitutes an objection to the Weak Phenomenal Principle. If, as Armstrong maintains, perception is nothing but the acquiring of knowledge or an inclination to believe, the notion of an immediate object of awareness is to be rejected even in the veridical case. In terms of the formal argument presented in Section 1.3, Armstrong's reductive analysis of perception presents an objection to our assertion of the Weak Phenomenal Principle at Premise 2.

Although, as Fish notes, few would now accept such an attempt to "reduce visual experience to the acquisition of belief" (2010, p. 66) it has suggested an "intentional theory of perception". Indeed, Byrne (2001), in the context of defending intentionalism, cites Armstrong as the "founding father of the view". Importantly, representationalist/intentionalist theories could be used to sustain a version of Armstrong's type-3 objection. Such objections are discussed in Section 2.2.3.

2.1.4 The time-gap argument

Having dismissed "The Argument from Causation" lightly (pp. 141-144), Armstrong goes on to examine "The Argument from the Time-Gap" at greater length (pp. 144-152), and affords it greater respect noting that "a complete emancipation from this argument is not very easy."

Here, in terms of our analysis in Section 1.3, Armstrong moves from the general argument from causation schema to focus on a particular instantiation thereof which draws on the standard causal picture to assert a particular fact F_n. However, although he introduces the argument from the time-gap as a "problem" which arises from "one special feature of the transaction between object and perceiver", his understanding of the relationship between "the argument from causation" and "the argument from the time-gap" is not clear.[56] Nevertheless his critique of the latter helps us understand both the version of direct realism he is defending and how his objections might be applied to the argument that we formalized in Section 1.3.3.

Armstrong opens his critique by contrasting Descartes' view of perception as instantaneous, with what "we know now", that is "in every case a time must elapse (sometimes a very short one) before a particular state of affairs is perceived." Thus, on the basis of these considerations, "one may be driven over once again to the view that the only thing that can be *immediately* perceived is something that occurs now, for example our *present* sense-impressions" and hence driven back to "some form of Representative theory". (p. 144)

---

[56] Whether or not Armstrong intends that the "argument from the time-gap" is to be understood as special case of the "argument from causation", is not clear. The two arguments are addressed in Part Four headed "The Argument from Causation", but this is in turn divided into two sections headed "1 The Argument from Causation" and "2 The Argument from the Time-Gap". Thus, confusingly, "The Argument from Causation" appears as both the chapter heading and as a section heading on the same level as "The Argument from the Time-Gap."

Armstrong's first response is to point out that in most cases the time-gap is very small and that consequently there is a "very good sense in which, in ordinary perception, the various events that are perceived, and the perceiving of them, all take place simultaneously." Since these ordinary perceptions are "indistinguishable from what they would be if light were propagated instantaneously", he concludes that "in most ordinary perception, there is no time-gap to worry about." Here is a type-2 irrelevance objection. However this is a only a preliminary and he focuses on those cases which, he accepts, do raise a difficulty for direct realism which must be addressed; that is where the time-gap is substantial, as in "seeing the sun or a star, hearing distant thunder." (p. 145)

The crux of his response to the difficult cases, against which the irrelevance objection cannot be deployed, is to challenge the "assumption that the time at which we perceive a certain state of affairs must be the same as the time at which that state of affairs obtains":

> So what we seem to have here [when we "look up at the sun] is a partially veridical perception of a state of affairs that occurred eight minutes before the perception of it occurred. *Prima facie* this falsifies the contention that we can only perceive what is going on *now*. **And if we accept the thesis that perception is the acquiring of knowledge of, or inclination to believe in, particular facts about the world by means of our senses** [bolded for reference below], it seems that there can be no objection to saying that when we look at the sun we acquire some knowledge of what was happening in a distant portion of space eight minutes ago. Since the distance is so great it is not surprising that the knowledge acquired is meagre, and that a large element to false belief or inclination to false belief is involved. (p. 146)

Armstrong takes the same approach with respect to the case of the astronomer observing the extinction of a star.[57] The astronomer acquires some knowledge of what was happening in a distant portion of space in the distant past, however "temporal illusion" is involved "because the star's extinction appears to be occurring now." (p. 148) Nevertheless Armstrong regards this as a case of immediate perception, since he allows that "immediate perception is sometimes a

---

[57] He considers briefly and rejects the suggestion that "what we immediately perceive is not the star but, but a *present* happening, causally connected with the extinction of the start many years ago. The star sends a message to us, as it were, and we immediately perceive the *message*, not the star." (1961, p. 147)

perception of past happenings." (p. 148) However, as he recognizes, this understanding of "immediate perception" is at odds with the principle that "what I now immediately perceive must exist now" and that he needs to explain why this seemingly uncontroversial principle is to be rejected. In drawing the chapter to a conclusion he notes:

> We have seen that in ordinary cases of immediate perception, in the paradigm cases where the objects we perceive are at no great spatial distance, the events perceived occur at the same time as the perception of them. … I suggest that this fact tends to be written into the concept of immediate perception. We tend to make it a logical necessity that what I now immediately perceive must exist now. There is a contingent fact involved, a fact about the speed of reaction of our sensory apparatus to physical events in our environment; but a conceptual necessity about immediate perception is erected on the foundation of this contingent fact. (p. 151)

This passage is highly significant with respect to our understanding of the version of direct realism which Armstrong is defending. There are two key points. Firstly, he not only rejects the claim that "what I now immediately perceive must exist now", but takes direct realism to be compatible with the contrary claim that what I now immediately perceive *need not* exist now. On this basis, Armstrong can conclude, as he does, that "the facts [with respect to the time-gap] involve no threat to our Direct Realism or our analysis of perception." (p. 152) However his reference to "our" direct realism and "our" analysis is important. Armstrong's acceptance of the above contrary claim, and hence his response to the argument from the time-gap, clearly depends on his reductive analysis of perception in terms of belief acquisition. Indeed he makes the dependency explicit on a number of occasions, as in the bolded section of the above quotation "**And if we accept the thesis that perception is….**" (p. 146) Here again, as in his rejection of "the argument from causation" his argumentation rests on his reductive analysis.

The second point to be drawn from the above quotation relates to the sense in which Armstrong should be understood when he advertises the fundamental question at issue in the monograph to be a "conceptual question" rather than a "scientific question". (p. xii Introduction) As we have noted, these introductory remarks appear to be at odds with his respect for, and lengthy

discussion of, the argument from science. However, here in his discussion of the time-gap, he is clearly addressing a question which is concerned with matters of "conceptual necessity" or "logical necessity" which are to be distinguished from matters of "contingent fact".[58]

Here in defending direct realism against the argument from the time-gap, Armstrong is claiming that when the notion of "immediate perception" is stripped down to the bare essentials of what is conceptually necessary such that "immediate perception" be the case, direct realism makes no claims which are incompatible with the finite speed of light or sound. This understanding is borne out in the continuation of the above quotation, where Armstrong fleshes-out the notion of immediate perception which his thesis of direct realism asserts:

> The discovery that we see the stars as they were many years ago comes as something of a shock to our conceptual system. But it does not force any *extensive* conceptual revision on us, for, when we look at the stars with this new knowledge, it still *looks* as if we are immediately perceiving present events. We need only say that what I immediately perceive must at least *seem* to exist now. (p. 151)

The implication is that the argument from causation would only go through if it forced an "*extensive*" conceptual revision on the direct realist, whereby he or she was forced to accept that "what I immediately perceive" does not "seem to exist now". We can certainly agree that the argument from the time-gap, or any other instantiation of the argument from causation schema, has no purchase against such a deflationary understanding of the thesis of direct realism which is concerned merely with what must "*seem* to exist". However this line of defence can offer no support to the thesis we are targeting, that is the metaphysical directness-claim *factually* construed which asserts, independently of how it seems, that the immediate object of awareness is one and the same as the physical object.

---

[58] We find a similar treatment of the time-gap argument in Suchting (1969, p. 50) which again suggests that a *strictly-conceptual* rather than a *factual* thesis is at issue. In response to the argument, he suggests that "the time of occurrence of the state said to be seen is only contingently related to the time of occurrence of event of it's being seen."

In summary, we take three points from Armstrong's time-gap section. Firstly, Armstrong is much less confident in his treatment of "the argument from the time-gap" than in his dismissal of "the argument from causation", noting with respect to the former that "Our argument has been complex and difficult, and in such a case error is only too easy." (p. 152) On our analysis, this difference is to be expected. In the latter, Armstrong is addressing an uninstantiated argument schema whereas in the former he has to contend with particular facts drawn from the standard causal picture, ie the finite speed of light and sound.

Secondly, as in his discussion of the argument from causation, Armstrong again draws heavily on his reductive analysis of perception in terms of belief acquisition, to underpin his argumentation at a number of key points. In our terminology, Armstrong is again raising a type-3 objection to the Weak Phenomenal Principle and the notion of an immediate object of awareness.

Thirdly, Armstrong's defence of direct realism, against the argument from the time-gap, is clearly a defence of a *strictly-conceptual* thesis which is defended on the basis of how things "must at least *seem*" if we are to avoid "any *extensive* conceptual revision" (p. 151). Moreover his discussion of the time-gap, suggests that his enigmatic "fundamental criticism" of the argument from causation should similarly be understood. It is becoming increasingly clear that, at least in his treatment of the argument from causation, Armstrong is not defending direct realism as a *factual* thesis. We draw the matter to a conclusion in Section 2.2.2.

2.2 More recent critiques

In comparison to the volume of literature devoted to illusion and hallucination, veridical perception and the argument from causation has received relatively little attention.[59] Indeed, as we have noted, a large proportion of the work over recent decades has been devoted to

---

[59] As Searle puts it, "A Martian Philosopher visiting Earth might be amazed at the amount of attention given to hallucinations in the philosophical discussion of perception and would not unreasonably conclude that hallucinations must be very common." (2015, p. 163)

developing and defending the representationalist and disjunctivist counters to the arguments from illusion and hallucination. Thus the philosophical literature focusing exclusively on causal arguments with respect to the veridical case is limited, and the material discussed below is taken from work which, as in Armstrong (1961), addresses the argument from causation in the context of a wider ranging examination of the traditional arguments.

This is not to say that issues relating to causation and the perceptual process do not appear in the literature. Causation comes to the fore when the principle of *same proximate cause same immediate effect* is used to defend the argument from hallucination against the disjunctivist's objection to the Common Factor Principle. Robinson calls this combination of causal considerations and the argument from hallucination "the revised – and successful – causal argument for sense-data". (1994, p. 151) Although disjunctivists ultimately reject this line of argument, it is not dismissed lightly and has received a good deal of attention. (Martin, 2004) (Nathan, 2005) (Byrn & Logue, 2008) (Fish, 2010, p. 89) However this argument is beyond the scope of our dissertation which is concerned only with the veridical case. In our terminology, Robinson's "revised causal argument" is a version of the argument from hallucination.

As we have seen, much of the literature is compromised by a failure to disambiguate the thesis of "direct realism" which is at issue. Moreover, when type-2 irrelevance objections are directed at the "argument from causation", they are often targeted on straw-men which attempt to infer their conclusions directly from the mere involvement of a complex causal chain of events, that is without reference to intermediate steps in the argument which draw on particular facts which arise from this the standard causal picture. Fortunately we do not need to resolve these issues. Our primary aim will be to assess the extent to which the objections tabled in the literature would, if so directed, raise objections to our formalized argument against the conjunction of the standard causal picture and the metaphysical directness-claim *factually* construed. Whether or not the authors intend to address this particular argument is secondary.

In our analysis of Armstrong's critique, we discussed the three types of objection that he raised in the context of the "argument from causation" and then considered the further issues arising in his response to the "time-gap argument" under a separate heading. Our discussion of the wider literature is organized similarly.

2.2.1 Type-1 objections revisited

As we have seen Armstrong presents but does not develop the objection that the conclusion of the argument, the rejection of direct realism, undermines the "evidence for the facts which are adduced as premises of the argument". (1961, p. 142)

This type of objection is only rarely found in the more recent literature. We suggest two reasons. Firstly, as in Armstrong, sophisticated objections may be seen as unnecessary since the argument is dismissed lightly on the basis of type-2 irrelevance objections. Secondly, type-1 objections to the argument from causation are closely related to the objections that were raised against indirect realism on the basis of the claim that it (indirect realism) leads inexorably to skepticism about the existence of an external world; this is in effect Berkeley's challenge to Lockean indirect realism (Armstrong, 1961, p. 30) (Dancy, 1985, p. 165). However it is now recognized that such objections depend on contentious claims about the epistemic consequences of direct realism, and assume that the rejection of direct realism would undermine natural science. Dancy puts the position well:

> This argument ["the sceptical objection"] has had an enormous effect, and has been the main impetus towards phenomenalism. …. For all its hoary past, however, there is a growing conviction that the argument is misconceived. If the argument makes crucial use of the idea that we cannot establish the required correlations because we cannot observe external objects, it is mistaken. …. And if the point is that indirect realism is vulnerable to the sceptical argument from error, the answer is that the argument from error is a menace to any theory of perception including direct realism and phenomenalism … (Dancy, 1985, p. 165)

Nevertheless, despite the "growing conviction" against the "sceptical argument", the self-refutation objection does, as Armstrong points out, have an "extra sting in its tail" (1961, p. 142)

when it is the validity of the argument from causation which is at issue. The traditional argument from causation relies on our perceptions as the basis for the standard causal picture on which it is grounded. Moreover, we will, in the following chapters, go beyond the standard causal picture to draw on the perceptually grounded claims which underpin the theories of perception which are posited by cognitive science. Indeed our examination of the isomorphisms involved in representation-friendly theories of perception will take us into what Dancy recognizes as the heart of the matter, that is the nature of the "correlations" between external objects and our perceptions thereof. Thus, whilst agreeing with Dancy's rejection of the "sceptical argument" for direct realism, we would nonetheless defend Armstrong's self-refutation objection as an objection to the traditional form of the argument from causation which takes the standard causal picture as a premise.

However such type-1 objections have no purchase against the conjunctive version of the argument developed in Chapter 1 which maintains only that the metaphysical directness-claim and the standard causal picture cannot both be true.[60] Similarly such objections have no purchase against the more sophisticated argument presented in Chapter 4, which maintains that the theories of perception posited by mainstream cognitive science are incompatible with the metaphysical directness-claim; here again our argument takes the form that both conjuncts cannot be true.

2.2.2 Type-2 objections revisited

Before addressing the irrelevance objections raised in the recent literature it is useful to examine the treatment of the argument in Locke (1967). Writing just a few years after Armstrong (1961), Locke's principal objection to the argument from causation appears to be similar to Armstrong's

---

[60] We grant that the traditional form of the argument ($N, N \rightarrow \neg D: \neg D$), which takes the standard causal picture, N, as a premise, is undermined by the objection that if the metaphysical directness-claim, D, is rejected we have no good reason for maintaining the standard causal picture. However, we do not take the standard causal picture as a premise, rather we argue only that it is incompatible with the metaphysical directness-claim ($N \rightarrow \neg D: \neg (D\&N)$).

enigmatic "fundamental criticism" and might therefore help to elucidate the latter. Locke

maintains that the argument "involves a serious misinterpretation of the scientific facts on which

it relies", that is "the familiar mistake of confusing perception with what is perceived". He notes

that this confusion is particularly obvious in Russell against whom he quotes Mundle (1959)

approvingly: "The argument ….confuses the *mechanism* on which the perception is dependent

with the *objects* that are perceived." (Locke, 1967, p. 116) In Locke's words:

> There are good reasons for identifying brain activity with perceiving …but none at all, that I can see, for identifying it with what is perceived. The fact that brain states may vary as what is perceived varies provides no reason for saying that the brain state is what is perceived …" (p. 117)

Here, as in the case of Armstrong's "fundamental criticism", if the argument were guilty as

charged it (the argument) should indeed be rejected as obviously confused. However the

argument from causation, as an argument against metaphysical direct realism, does not assume

or assert that "the brain state is what is perceived". Again, as in Armstrong, the question is why

Locke should suppose that the argument involves such a fallacious identity claim. One answer is

that Locke takes the argument to be, not merely a negative argument against direct realism, but a

positive argument for a particular alternative theory: that is a version of indirect realism which

maintains that brain states are the perceptual intermediaries. Against such a positive argument,

Locke's objection holds, but it has no purchase against the purely negative argument we present.

In what follows, we will find similar loosely targeted objections in the recent literature.

However Locke raises a type-2 irrelevance objection which could be directed at any argument

which draws on facts concerning the physical and physiological mechanism of perception. In his

introduction, he discusses the type of question that "philosophical theories of perception" are

best thought of as giving answers to and concludes:

> What sort of things do we perceive, and how are they related to the sorts of things that we ordinarily think of ourselves as perceiving. This question may not look philosophical, but

> it becomes philosophical in as far as it is, and can only be, answered not by reference to any empirical facts, but some form of theoretical argument. (1967, p. 14)

If such a question is at issue, then all considerations with respect to the physical and physiological mechanisms of perception are indeed irrelevant. We dub this the "overarching irrelevance objection". Clearly Locke is defending direct realism construed as a *strictly-conceptual* thesis which places a restriction on the admissible data which may count for or against, and thereby excludes the very facts on which the argument from causation depends. He is not defending the *factual* thesis we are targeting.

We found a similar passage in Armstrong's introduction (1961, p. xii). However, whereas Locke is unequivocal with regard to the irrelevance of "any empirical facts", Armstrong's position was less clear. Nevertheless, on balance, drawing particularly on his otherwise enigmatic "fundamental criticism", and his response to the argument from the time-gap, it seems we must understand Armstrong, similarly, as defending a *strictly-conceptual* rather than a *factual* thesis. Perhaps we need to take his introductory remarks, where he advertises the question at issue as a "conceptual question", at face value.

The position defended by Locke (explicitly) and Armstrong (implicitly) appears to be that which Ward describes in his examination of the direct realist's response to the argument from causation:

> … the causal processes are merely a *physical precondition* of direct perception, i.e., are merely processes which, contingently, must always occur if the observer is to be able to perceive objects directly. It is admitted, *as a matter of fact*, an observer would not be able to perceive an external object unless the causal processes produced his awareness of the object. But it is not usually admitted that our grounds for deciding whether someone can be directly perceiving an external object need make any reference to the causal processes. … the causal processes of perception need not enter into the **logical analysis of perception** [emphasis added] at all. (Ward, 1976)

We conclude that, contrary to our provisional assumption in Section 2.1, when Armstrong rejects the argument from causation "out of hand" (1961, p. 144), he is not defending the *factual* thesis

we are targeting. Rather he is defending direct realism as a *strictly-conceptual* thesis, that is a particular "logical analysis of perception", with respect to which the facts adduced by the argument from causation are inadmissible. From this perspective Armstrong's seemingly enigmatic "fundamental criticism" can be explained. However, it has no force against our arguments.

Most of the recent literature, as in Armstrong's "fundamental criticism" (1961, p. 142), and Locke's "main fault" (1967, p. 116), continues to give primacy to versions of the irrelevance objection. However, the objection is now very rarely presented explicitly in the context of an overarching claim that empirical observations and, a fortiori, the theories of cognitive science, are as a matter of principle irrelevant to the philosophical thesis of direct realism. Indeed such a demarcation between the philosophy and science of perception, explicit in Price (1950, p. 2) and Ryle (1953, p. 185), is now rarely asserted and would be very difficult to sustain in the light of developments discussed in Chapter 3.

Now, it is not empirical evidence in general which is deemed to be irrelevant but, rather, the particular evidence which is adduced by the particular version of the argument which is tabled for examination. The question is whether, in the absence of the overarching irrelevance objection, any of these particular irrelevance objections have force against our formalized argument. We maintain that they do not.

Dancy presents a typically brief dismissal of the argument from causation in his discussion of indirect realism. As is often the case, he alludes to, rather than states, the argument:

> Given that complexity [of the causal processes involved in perception], it asks, how can we claim that we perceive the external object directly? There are many states or processes of the brain intermediate between the external object and the perception; surely then the object is separated from us, and can only be perceived indirectly by means of the effects it has on our retinal surfaces etc. (1985, p. 154)

Although this certainly captures the spirit of the case against direct perception, it falls far short of

a rigorous argument. The connection between the premise, that a complex process is involved,

and the conclusion, that perception is indirect, is claimed on the basis that it would be

implausible to suppose otherwise, rather than deduced. However Dancy does not challenge the

argument on this point. Rather, his objection is that such neurophysiological evidence is

irrelevant, since it does not establish that perception is indirect in the sense that would be

required of a successful argument for indirect realism:

> But, crucially, we are not aware of their [brain processes'] occurrence in any sense even
> distantly analogous to the sense in which we are aware of external objects. So the
> neurophysiological processes do not function as intermediary direct objects of perception.
> (1985, p. 154)

We do not need to take issue with Dancy, though there is an important point to be taken from his

treatment of the argument. He is rejecting the argument from causation as a positive argument

for indirect realism; in particular he is targeting indirect realists who might maintain, on the basis

of the argument, that brain processes are the immediate objects of awareness. However, we will

need to take issue when, as below, the same irrelevance objection is used against the argument

from causation presented as a purely negative argument against the metaphysical directness-

claim. The importance of the distinction between a positive and negative argument with respect

to the nature of perception is emphasized in Section 1.2.4.

Searle addresses the argument from causation[61] as part of a defence of his "default positions"

(1998, p. 9) [62] which include a direct realist understanding of perception. Here again the

argument is alluded to, rather than stated:

> If you consider scientifically what happens when you see a tree, here is what you find:
> Photons are reflected off the surface of the tree, they attack the photoreceptor cells in the

---

[61] He refers to it as the "argument from science". (Searle, 1998, p. 28)

[62] "Default positions are the views that we hold prereflectively so that any departures from them requires conscious
effort and a convincing argument." (Searle, 1998, p. 9)

retina ….eventually this series of neuron firings causes a visual experience somewhere deep in the brain. All that we see, literally, is the visual experience in our brains. …. the basic idea is that perceivers don't actually see the real world. (Searle, 1998, p. 28)

Searle deploys a version of the irrelevance objection. Now it is the fact that "I can give a causal account", that is irrelevant to the standing of direct realism:

From the fact that I can give a causal account of how it comes about that I can see the real world, it doesn't follow that I don't see the real world. It [the argument] is, indeed, a variant of the genetic fallacy. The fact that I can give a causal account of why I believe that two plus two equals four (I was conditioned by Miss Masters, my first grade teacher) does not show that two plus two does not equal four. (1998, p. 28)

On this point we can agree. If the argument did depend on such an implication it would obviously be fallacious. Indeed, it is not clear why Searle presents such a straw man[63] for examination; he does not give any examples of the argument actually being presented in this form. However he continues, seemingly on the basis of rejecting this somewhat spurious version of the argument, to claim that:

There is no inconsistency between asserting, on the one hand, "I directly perceive the tree", and asserting, on the other, "There is a sequence of physical and neurobiological events that eventually produce in me the experience I describe as seeing the tree." (1998, p. 29)

Here we do take issue with Searle. Although the argument under examination has not shown there to be any "inconsistency", we maintain that our formalized argument does show there to be an inconsistency when "directly perceive[ing] the tree" is taken to imply the metaphysical directness-claim *factually* construed. Whether or not Searle is defending this claim is immaterial[64], the key point is that his critique fails to address our formalized argument. Indeed there is no point of contact between Searle's critique and our argument. (As discussed in Section 2.2.3, Searle presents a type-3 objection in his recent monograph (2015).)

---

[63] McDermid suggests that Searle has "made things too easy for himself by focusing exclusively on a rather weak version of the argument." (2004)

[64] Searle's response to the argument in terms of his reference to his belief that two plus two equals four suggests that his is taking it to be an argument against epistemological direct realism and that, at this point, he is defending the epistemological directness-claim. However the main thrust of the monograph suggests that he would also wish to defend the metaphysical directness-claim factually construed. See particularly his "default positions." (Searle, 1998, p. 9)

Written four decades after Armstrong (1961), Huemer's *Skepticism and the Veil of Perception* (2001) is one of the few recent monographs which present a defence of direct realism against all the traditional arguments. However, as does Armstrong, he dismisses the argument lightly. Again, the argument is regarded, not merely as invalid, but patently so.

Huemer outlines the "worry" which he takes to be at the core of the argument. Here again we have an allusion to, rather than a statement of, the argument:

> In order for me to see the cup on the table, for instance, light rays have to travel the distance between the cup and the eye. Then electrical signals have to travel down my optic nerve. Then my brain has to process the information. All these events intervene between the object I take myself to be seeing and my experience of seeming to see it. Therefore, must we conclude that I am not directly aware of the object? (2001, p. 135)

He then goes on to show that our worries are based on a misunderstanding:

> This sort of worry rests upon a misunderstanding of the term "direct awareness". If being directly aware of a thing means being directly aware of it with no causal intermediaries between it and your state of awareness, then direct realism would be swiftly dispensed with on empirical grounds. But if being directly aware of a thing means having awareness of it not based on one's awareness of anything else, then these considerations are irrelevant, for the aforementioned processes intervening between the cup and my experience of the cup do not include any states of awareness. They are blind physical processes with a mental state at the end. … Brain processes cause my visual experience, but I am not seeing brain processes; I am seeing the cup. (2001, p. 135)

Huemer's version of the irrelevance objection is that the considerations with respect to the "causal intermediaries", say brain processes, are irrelevant since we have no "awareness" of such intermediaries.

According to Huemer, the argument rests on a misunderstanding about what direct realists mean by "being directly aware of a thing". The implication is that proponents of the argument from causation are taking direct realists to assert more than they actually do. We agree that direct realists do not deny the involvement of causal intermediaries; indeed, as discussed above, they

accept the standard causal picture. However, Huemer reduces direct realism to a purely negative claim, which could only be compromised by the argument if it demonstrated that we had an awareness of such intermediaries. Indeed Huemer's concluding sentence implies that the argument from causation would only be successful if it could be shown that we are aware of "brain processes" in the sense that we are aware of "the cup". On this understanding, it is not surprising that he can dismiss the argument so easily. However, as a purely negative argument against the metaphysical directness-claim, the argument from causation does not, and need not, make such spurious claims.

We grant that if the argument from causation was presented as establishing an indirect realist theory according to which we are, in the above sense, aware of the causal intermediaries in the perceptual process, it should certainly be rejected. However Huemer's objection cannot serve in the defence of direct realism and, a fortiori, has no point of contact with our formalized argument against the conjunction of the metaphysical directness-claim and the standard causal picture.

We conclude with le Morvan's frequently cited paper *Arguments Against Direct Realism and How to Counter Them* (le Morvan, 2004) Here he addresses the "eight main arguments"[65] against direct realism and gives more or less equal weight to each. His aim is stated as the "exploration of strategies Direct Realists may deploy against arguments purporting to defeat Direct Realism". His treatment of the argument from causation[66] serves as an exemplar of a plausible objection to a loosely formulated argument targeted on a loosely defined thesis, which nevertheless fails to address our formalized argument against metaphysical directness-claim *factually* construed.

---

[65] "The Causal Argument, the Time-Lag Argument, the Partial Character of Perception Argument, the Perceptual Relativity Argument, the Argument from Perceptual Illusion, the Argument from Hallucination, the Dubitability Argument, and the Objective Feature Argument"

[66] Le Morvan uses the term "causal argument".

Unfortunately, he makes no mention of the distinction between epistemological and metaphysical direct realism. This is a serious omission in the context of his objectives, which fails to recognize that at least some objections, to some of the eight arguments, might require separate consideration in the context of epistemological and metaphysical directness-claims. Indeed we maintain that the argument from causation is a case in point: it is successful as an argument against the metaphysical directness-claim, but does not have the resources to address the epistemological directness-claim.

Le Morvan presents the argument from causation thus:

> First Premise. Direct Realists hold that external physical objects or events can be immediate or direct objects of perception.
>
> Second Premise. But perception involves a long and complex causal series of events. For instance, light quanta are reflected or emitted from an external object, the light quanta then travel through an intervening medium (e.g., air and/or water), they then hyperpolarize retinal cells by bleaching rhodopsin photopigment molecules, and then a very complex series of physiological processes takes place in the eye and in the brain eventuating in perception.
>
> Conclusion: Direct Realism is false. Given this long and complex causal series, physical objects or events cannot be immediate or direct objects of perception. (2004, p. 222)

The "First Premise" is a definition of direct realism, which appears to align with our definition of the metaphysical directness-claim. Here we take le Morvan to hold that, in the standard veridical case, "external physical objects or events" are one and the same as the "immediate or direct objects of perception". The "Second Premise" is an assertion of the standard causal picture fleshed-out in terms of contemporary physics and physiology. The "Conclusion" is the rejection of direct realism as defined in the "First Premise". However here again, despite the formal layout, we have an allusion to, rather than a statement of, the argument; there is no attempt to show how the conclusion is deduced from the premises.

Le Morvan's response to the argument is based on the distinction between two claims:

(i) perception is indirect in the sense that it involves a series of causal intermediaries between the external object or event and the percipient:

and

(ii) perception is indirect in the sense of involving a prior awareness of something other than the external object (or event).

On the basis of this distinction he maintains:

> What the Causal Argument establishes is only the *causal* indirectness of perception in the sense of (i), not the *cognitive* indirectness in the sense of (ii). Hence, this argument does not refute Direct Realists not committed to denying the indirectness of perception in the sense of (i). Thus, lest they fall prey to this argument, Direct Realists should be careful to distinguish between causal indirectness and cognitive indirectness. (2004)

Here we have a version of the irrelevance objection: the causal considerations adduced in the "Second Premise" are irrelevant to the standing of direct realism since they fail to establish claim (ii), with the implicit corollary that these considerations would only be relevant if they did establish claim (ii). This is, in effect, a more formally stated version of the objection that we found in Huemer. However it has no purchase against our version of the argument from causation. There are three key points.

Firstly, we agree that the argument, as le Morvan presents it, does not show that (ii) follows from (i), indeed it fails to establish any substantial conclusion. Here, in common with all the versions so far tabled for examination, the argument omits the crucial intermediate stages between the assertion of the "Second Premise" and the "Conclusion". The argument is presented as if the conclusion, the falsity of direct realism follows immediately from the premise asserting the standard causal picture. However there is no such immediate connection, and the intermediate stages of the argument, along with the auxiliary premises, need to be presented. Such is the argument developed in Section 1.3.

Secondly, contrary to le Morvan's analysis, a successful argument against direct realism does not need to establish claim (ii). Claim (ii) is an assertion of indirect realism, which goes beyond the rejection of direct realism. A successful argument only needs to demonstrate the falsity of the

direct realist theory of perception, say, as defined in "First Premise", and need make no positive claim, say, as asserted in claim (ii).

Thirdly, we maintain that our argument against the metaphysical directness-claim, which includes the necessary intermediate stages and auxiliary premises which are absent from le Morvan's straw man, constitutes a successful argument against direct realism as defined in his "First Premise". The fact that our formalized argument does not establish claim (ii) does not constitute an objection. Le Morvan's critique has no force against our formalized argument and leaves it unscathed.

2.2.3 Type-3 objections revisited

As we have seen, Armstrong's analysis of perception in terms of belief acquisition is the mainstay of his defence of direct realism against the arguments from illusion and hallucination, and also underpins his critique of the argument from causation at a number of key points. Indeed we accept that Armstrong's rejection of the Weak Phenomenal Principle, if it were sustained, would constitute an objection to Premise 2 of our formalized argument.

However as we noted in Section 1.1.2 the Weak Phenomenal Principle is a difficult target to attack and most contemporary direct realists would be unwilling to follow Armstrong in rejecting the claim that there is, in the veridical case, something of which the subject is aware. Indeed, citing both Armstrong and Dennett as philosophers who maintain "that *all* perceptual consciousness, veridical as well as illusory, should be construed intentionally", Smith notes:

> On such a view of perception, all one is allowed to attribute to a conscious subject on the cognitive side of things are thoughts, beliefs or dispositions towards behaviourally specified discriminations, or "representations". We have here a recrudescence of the "feigned anaesthesia" with which the early behaviourists were commonly taunted. (2002, p. 46) [67]

---

[67] That thoroughgoing reductive intentionalism should have these highly counter intuitive implications is not surprising when we recall (Section 2.1), as Smith (p. 45) reminds us, that Armstrong, arguably the founding father of the position, "proposed to account for all forms of perceptual awareness as being simply a matter of acquiring

However, though Smith rejects this account outright, he admits that "Such a reductive, or eliminative, account of sensory experience is astonishingly widespread today." Indeed so, and we will accept that some versions of contemporary representationalism provide a basis for an objection to the Weak Phenomenal Principle at Premise 2, but first it is useful to consider the strength of the opposition.

Disjunctivist direct realists divide with respect to the (strong) Phenomenal Principle; negative disjunctivist, eg Martin, (2004), deny that there is an object of awareness in the hallucinatory case. However they would endorse the Weak Phenomenal Principle[68] and affirm that there is an object of awareness in the veridical case. (Though, as discussed below, they might nevertheless object to our call on the notion of "immediate object of awareness" in the formalized argument.) Indeed Martin suggests that both the sense-datum and representationalist theories of perception alike "amount to error-theories of sense experience" which are at odds with "a natural conception of how our veridical experience relates us to the world around us." (2004, 84) Similarly, Crane cites McDowell as a direct realist who argues that the representationalist theory of perception is "no better than a sense-data theory in being able to do justice to the intuition of openness [to the world]." (Crane, 2005)

Although representationalist direct realists reject the (strong) Phenomenal Principle, they divide with respect to the Weak Phenomenal Principle and the notion of an immediate object of awareness in the veridical case. In the first camp, representationalists would accept that there are immediate objects of awareness in the veridical case, indeed they are one and the same as the

---

beliefs (of a specified sort)about the environment." Earlier Nelson (1964, p. 160) and Locke (1967, p. 29) take issue with Armstrong on this point.

[68] As Fish puts it in his discussion of disjunctive theories: "Typically, the motivation for endorsing disjunctivism about phenomenology is to defend a philosophical theory about the good cases of perception known as naïve realism or relationalism. … A useful way of understanding this claim is to see the naïve realist as endorsing the Phenomenal Principle, at least for the limited case of perception." (2010, p. 96) Fish uses the term"perception" to mean veridical perception as opposed to cases of illusion or hallucination. (p. 2)

physical objects perceived. Crane outlines such a version of representationalism/intentionalism in his Stanford Encyclopedia entry on *The Problem of Perception*:

> What are the intentional objects of perceptual experience, according to intentionalists? In the case of veridical perception, the answer is simple: ordinary, mind-independent objects like the churchyard, the snow (etc.) and their properties. (Crane & French, 2015)

Smith would agree. He holds that the immediate objects of awareness are intentional objects which, although in the case of hallucination they (the intentional objects) are "nonentities" or "non-existents" of questionable ontological status[69], they are straightforwardly physical objects in the veridical case. (2002, pp. 234, 257) The question thus arises as to how Smith might defend direct realism against our argument from causation. Significantly, his substantial monograph focuses exclusively on the arguments from illusion and hallucination, and he accepts that direct realism might be shown to be false for other reasons.[70] (2002, p. 5)

However, in the second camp, thoroughgoing reductive representationalists follow Armstrong all the way in rejecting the Weak Phenomenal Principle, and would thus reject the assertion thereof at Premise 2 of our formalized argument. Clearly, if sustained, this would constitute a conclusive objection. We will take the likes of Dennett (1993) and proponents of "Strong Representationalism", the thesis that "phenomenal character is one and the same as representational content that meets certain conditions" (Tye, 2009), when reductively construed, to be within the Armstrong-camp.

Not surprisingly, since both disjunctivists and Smith-camp representationalists accept the notion of an immediate object of awareness in the veridical case, objections to the Weak Phenomenal Principle are rarely raised in discussions of the argument from causation; only Armstrong-camp

---

[69] Smith notes that "intentional objects are, for many people [analytical philosophers], creatures of darkness." (2002, 237)

[70] "Kant's arguments about the ideality of space ….would constitute, if sound, a refutation of direct realism." And following reference to Husserl, Robinson and Foster's position contra direct realism: "So perhaps Direct Realism is false. What this work is intended to show is that if this is indeed the case, it cannot be shown to be so by reflecting on the nature of perception." (Smith, 2002, p. 5)

thoroughgoing reductive representationalist would be in a position to do so. Moreover, since the argument from causation is often dismissed lightly on the basis of type-2 irrelevance objections, issues concerning the Phenomenal Principle, which are central to the discussion of the arguments from illusion and hallucination, are rarely addressed.

Robinson (1994) is an exception. Moreover, he is unusual in that he presents a version of the argument for examination which both disambiguates its target and includes the intermediary stages in the argumentation. Robinson's *Perception* (1994) again addresses the full range of arguments against direct realism though, in contrast to the other authors discussed herein, he proposes a sense datum theory of perception.[71] After some discussion and refinement, Robinson presents the argument from causation as follows:[72]

1. Perceptual experience occurs at the end of a causal chain running from the external object to the brain.
2. The content of perceptual experience is part of, or internal to, the experience.
3. If something occurs at a certain location anything which is part of or internal to it occurs at (or within) that location.
   Therefore
4. The content of perceptual experience occurs at the end of a causal chain running from the external object to the brain.
   The argument is then completed ….
5. Things located at opposite ends of a causal chain cannot be identified.
   Therefore
6. The content of perceptual experience is not identical with the external object perceived. (1994, 86)

Here, there are parallels with our formalized argument.[73] Robinson's "external object" and "content of perceptual experience" correspond, respectively, to our "normal physical object" and "immediate object of awareness". In our terminology, stage 4 is the introduction of a "particular

---

[71] He concludes that naïve realism has been "conclusively refuted", accepts some form of sense datum theory and suggests a preference for phenomenalism over representational realism. (1994, 238)

[72] The text is Robinson's but the steps have been renumbered.

[73] At this point in the discussion we can ignore the distinction between the conjunctive and non-conjunctive versions.

fact" F_n, the assertion of which rests on both the Weak Phenomenal Principle **P** and the naturalistic standard causal picture **N**. Stage 5 draws on the indiscernibility of identicals **I** to show that, contrary to the metaphysical directness-claim, the normal physical object and the immediate object of awareness cannot be one and the same.

Presented in this way, the argument avoids type-2 irrelevance objections on two counts. Firstly there is no ambiguity in the thesis which is being targeted. The argument is presented as a purely negative argument against the metaphysical directness-claim; there is no suggestion that the argument purports to, or needs to, establish a case for indirect realism. Secondly, the relevance of the causal considerations, with respect to the thesis at issue, is clearly demonstrated. Robinson's argument draws on a particular fact arising from the standard causal picture to establish the discernibility of, and thus non-identity of, the "content of perceptual experience" and the "external object", rather than merely alluding to the supposed implausibility of direct perception given the complexity of the perceptual process.

However Robinson does not accept the argument as it stands. His concern, as he notes at the outset, is that the argument must establish "that *what I am aware of* is not the external object, rather than *my being aware of it* is not something external", and that the second interpretation "would be harmless" to direct realism. (1994, p. 85) In terms of his formal argument, he is concerned that the notion of the "content of perceptual experience" may not be strong enough to sustain his stage 3. In particular he notes that there "is a weak sense of 'content' …. which makes it synonymous with 'intentional object'" and that on this interpretation[74] it would be incorrect to claim that the "experience and its content cannot be separated" and must therefore be co-located. For this reason, his conclusion is that, as "so far stated", the argument from causation

---

[74] "It could be argued that Charles I is internal to or part of my thought that Charles I was a martyr, but that he does not occur at the same location as my thought but 340 years prior to it (or alternatively, in heaven whilst my thought is mundane)." (Robinson, 1994, p. 87)

"has little force." (1994, p. 89) Here, Robinson is raising the representationalist objection to the Phenomenal Principle. However the objection is soon to be rejected.

Since Robinson has no special interest in arguments with respect to the veridical case alone, he does not consider the argument from causation further. Rather, he develops the "revised causal argument" which combines the argument from causation and the argument from hallucination to defeat the disjunctivist objection to his case for sense data.[75] (Robinson, 1994, p. 151) However he goes on to reject the intentionalist understanding of the content of perceptual experience and to present a substantial defence of the Phenomenal Principle. (1994, p. 163) Later he concludes:

> As the great philosophers of the early part of this century thought, the Phenomenal Principle is sound, and if I clearly see something red then there is something red of which I am aware. (1994, p. 186)

Thus, in terms of our formalized argument, Robinson is raising, and then rejecting, the representationalist's type-3 objection to the assertion of the Weak Phenomenal Principle at Premise 2.

We agree with Robinson's analysis of the argument. When the target is disambiguated as the metaphysical directness-claim *factually* construed, and the argument is clearly stated, the various irrelevance objections on the basis of which it has been dismissed lightly do not arise. Only type-3 objections remain to be addressed.

However, although, only the Armstrong-camp representationalists would reject the Weak Phenomenal Principle (Premise 2), others might object to our identification of the "something of which the subject is aware", as in the Principle, with the "immediate object of awareness" as it is used in the argument (Premises 6 and 7). Indeed, as discussed in Section 1.2, notwithstanding our explication of the immediate object of awareness as the bearer of phenomenal character, opinions divide and some (some non-reductive intentionalists and disjunctivists) would maintain that phenomenal character is born by the perceptual experience itself rather than by some

---

[75] He advertises the argument as "The Revised – and Successful - Causal Argument for Sense-data" (Robinson, 1994, p. 151)

immediate object of awareness. From this perspective, it is possible to accept the Weak Phenomenal Principle and the notion of something of which the subject is aware, thus avoiding Smith's charge of feigned anaesthesia, whilst refusing to sanction the arguments call on the notion of an immediate object of awareness to assert the particular fact F_n.

Moreover, direct realism is sometimes defined purely in terms of experiences without any reference to immediate objects of perception. For example Fish introduces the thesis in terms of the "claims [that] the naïve realist makes about the nature of the experiences we enjoy when we [veridically] perceive the world":

> As far as the conscious character of visual experience is concerned, the core claim of naïve realism is that, when we see, external objects and their properties "shape the contours of the subject's conscious experience" (Martin 2004: 64), where the metaphor of 'shaping' is read in a constitutive rather than a mere causal sense … they shape the contours of the subject's conscious experience by actually being the contours of the subject's conscious experience. (Fish, 2009, p. 6)

Similarly Logue introduces the thesis in her paper *Why Naive Realism?* thus:

> Naive realism is the view that veridical experience fundamentally consists in the subject perceiving things in her environment and some of their properties … [That is] veridical experience *fundamentally* consists in such a state of affairs … [and] this state of affairs constitutes the *metaphysical structure* of veridical experience. (Logue, 2012, p. 211)

We maintain that the notion of an immediate object of awareness is essential to the thesis of (metaphysical) direct realism, not only as it is defined by Smith, but also as it has been discussed since the Early Moderns, largely rejected until the middle of the last century, and now defended within contemporary analytic philosophy. That is we understand metaphysical direct realism to be, in essence, an identity claim which aligns with our pre critical intuition that "these direct objects of perception are the persisting mind-independent objects we all know and love." (Brewer, 2007, p. 87) On this understanding we might argue that our interpretation of the "something of which the subject is aware" as an immediate object of awareness (rather than merely a perceptual experience) can only be avoided at the expense of a deflationary and inadequate construal of the thesis which is being defended.

However we will not argue on this basis. It would be, to say the least, contentious to defend our position on the basis that our opponents are not really "direct realists" after all. We will grant that direct realist can accept the Weak Phenomenal Principle whilst rejecting the notion of an immediate object of awareness as it is used in the formalized argument. Thus although we maintain that the argument is robust, it is open to objections not only from those who reject the

Weak Phenomenal Principle outright, but also those who reject our interpretation thereof. Such an objection might be found in Searle's recent defence of direct realism, *Seeing Things As They Are*. (2015)

Although, as we have seen in Section 2.2.2, Searle has voiced type-2 irrelevance objections, he presents a type-3 objection in this recent monograph. Unfortunately, in common with much of the literature, he gives scant attention to the argument from causation and focuses on the arguments from illusion and hallucination. However he believes that all three arguments rely on the same fallacy (2015, p. 22); that is, they rest on "The Bad Argument" which has led philosophers astray[76] from the Early Moderns through to the present[77]. Though, significantly from our perspective, he admits that reliance on the fallacy is "not so obvious" in the case of the argument from causation.[78] (p. 28)

Searle would not reject the Weak Phenomenal Principle outright and deny that there is, in the veridical case, something of which the subject is aware. Indeed, citing Dennett and Campbell as proponents thereof he rejects the view that, put bluntly[79], "perceptual conscious does not exist at all", as "so implausible it is hard to imagine anyone ever defending it." (p. 49) However Searle would object to the call on the notion of an "immediate object of awareness" at Premise 6 and 7 of our formalised argument. He would maintain that our argument, no less than the others, rests on an ambiguity in the sense of "aware of". According to Searle:

> …the intentionality of the visual experience enables the experience to be an experience of objects and states of affairs in the world; but in the sense of "aware of," in which the subject is aware of the experience, the experience and the awareness are identical. …
> And one can be both "aware of" (in the constitutive sense) the experience and "aware of"

---

[76] Against this claim we would endorse Martin's analysis. After a discussion of the treatment of intentionality by the early moderns up to Russell, Broad and Price, he suggests that the evidence does not support the conclusion that they suffered from " a general confusion about the notion of intentionality which explains their endorsement of the argument from illusion…" (2000, p. 19)

[77] Searle believes that "the Disjunctivists" accept "The Bad Argument". (2015, p. 29)

[78] Searle uses the term "Argument from Science" for the argument which draws on the mechanism of perception.

[79] As Searle notes, it is rarely put so bluntly.

(in the intentionality sense) the ontologically objective objects and states of affairs in the world that the experience presents as its intentional conditions of satisfaction. (2015, p. 28)

Searle's notion of intentionality is difficult to understand[80] and it is unclear, exactly, how his charge of reliance on "The Bad Argument", relates to the veridical case and our argument from causation. However we will take Searle to be an influential contemporary philosopher who, whilst accepting the Weak Phenomenal Principle, would object to our call on the notion of an immediate object of awareness. Such objections remain to be addressed along with those from the Armstrong-camp who would reject the Weak Phenomenal Principle outright.

2.2.4 The time-gap argument revisited

Before concluding this chapter, we return to the argument from the time-gap which, on our analysis, is a particular instantiation of the argument from causation schema. As such it provides a test-case which requires a more focused response from its critics than do the more loosely formulated arguments, that is un-instantiated schemas, which are often tabled as "the argument from causation". Moreover, as we saw in Armstrong (1961), the manner in which the time-gap version of the argument is treated helps us to understand the nature and content of the direct realist thesis which is being defended.

The main criticism of the argument from causation, type-2 objections that the considerations adduced by the argument are irrelevant to the thesis at issue, is rarely directed at the argument from the time-gap. It is generally accepted that, at least in the case of the disappearing star, direct realism has a prima facie case to answer. The primary response is, as in Armstrong, to challenge the assumption that "the time at which we perceive a certain state of affairs must be the same as

---

[80] For example, Travis suggests, in a critical review, that Searle's view "comes [close] to that idea, sense data, on which he [Searle] heaps abuse." (2005)

the time at which that state of affairs obtains." (1961, p. 147) Locke, Dancy, Huemer and le Morvan all focus on the case of the disappearing star and provide a similar response.

Locke addresses the argument briefly and notes that "such examples show that our naïve common sense beliefs about what we perceive will have to be altered in some way" and suggests that "we could even say" that "we can see 'into the past', i.e. see what no longer exists". (1967, p. 118) Similarly, Dancy maintains that it is wrong to suppose that "the direct object of perception must be present at the moment of perception" (1985, p. 152), and le Morvan rejects the claim that "If something no longer exists, we cannot perceive it". (2004, p. 223)

Huemer devotes a relatively lengthy section of his monograph to the "time-gap argument". However, he concedes that his response to the case of a star which was "destroyed 300 years ago" may "seem disappointingly simple":

> What are you aware of when you are looking at the star 1000 light years away? You are aware of the star, as it was 1000 years ago. I see no reason why one should not be able to perceive something in the past. …if you can hear an event after the time at which it happened, why shouldn't you be able to see an object after the time at which it existed? (2001, p. 132)

And, similarly, in his more recent *Stanford Encyclopedia* entry on "Sense Data":

> The natural reply for a theorist wishing to resist sense data is to claim that one can "see into the past", that is, that one's perceptual experiences may represent past states of affairs, or represent objects as they were at an earlier time. (Huemer, 2011)

In summary then, in the case of the disappearing star, the direct realist's response is to maintain that we are seeing an object which no longer exists.[81]. Moreover the implication is that this response is compatible with the direct realist thesis which is being defended. Le Morvan puts it explicitly:

---

[81] The alternative response, that we are seeing the light emitted by the star rather than the star itself, is sometimes discussed only to be rejected. (Armstrong, 1961, p. 147) (Huemer, 2001, p. 132) Clearly it would raise difficulties for direct realism.

> It's important to keep in mind here that Direct Realists need not be committed to the claim that we can *now* be aware of the no-longer existent object as it is *now*, but only that we can *now* be aware of the once-existent object as *it used to be*. (le Morvan, 2004)

We agree that, if the epistemological directness-claim with respect to the role of inference, or some *strictly-conceptual* claim, is at issue, there is no obvious incompatibility and the argument from the time-gap fails. However, when it is the metaphysical directness-claim which is at issue, that is the identity of the normal physical object and the immediate object of awareness, the response that we can be perceptually "aware of the no-longer existent" object, albeit as "it used to be", can offer no support. Indeed such a response would appear to be a capitulation to the argument. Granting the argument's call on the notion of an immediate object of awareness, which is not at issue in the four critiques cited above, the implication of this response is that, contrary to the metaphysical directness-claim, the immediate object of awareness exists when the normal physical object does not.

It might be objected that our response to le Morvan et al is somewhat cavalier and that we have given no good reason to justify our assertion that, if direct realism is true, the physical object has to exist at the time of its perception. As noted above, "direct realism" can be construed (as an epistemological or *strictly-conceptual* thesis) such that there is no such requirement; the requirement is peculiar to metaphysical direct realism *factually* construed. The crucial point is that the metaphysical directness-claim is an identity claim: the immediate object of awareness **A** in standard perceptual situations is (one-and-the-same-as) a normal physical object **B**. Thus, if **A** and **B** are identical it cannot be the case that **A** exists at a time that **B** does not exist. In terms of Smith's taxonomy of the arguments against direct realism, it draws on the type-c fact that "one exists at a time when the other does not." (2002, p. 8)

Robinson presents a similar defence of the "time-lag argument" with respect to the disappearing star; he dubs this the "restricted version" and recognizes that the generalized version involves further argumentation and is less "intuitively obvious." (1994, p. 80) In Robinson's statement of

the argument, he captures the crucial point with respect to identity thus: "If one thing exists at a time when the other does not then they are not identical". Moreover he notes, "[this premise is] an uncontroversial application of Leibnitz's Law, which says that *a* and *b* cannot be identical if one posses a property – in this case existence at a given time – that the other lacks." (p. 81) Here, as in his discussion of the "causal argument", Section 2.2.3, Robinson avoids the irrelevance objection by clearly targeting the metaphysical-directness claim.

We have noted at a number of points that critiques of the argument from causation are often compromised by a failure to disambiguate the thesis which is at issue, and the case of the disappearing star is a particularly clear example of the problems which arise. It goes unrecognized that a response which might serve in defence of the epistemological directness-claim not only fails in the defence of, but constitutes a serious attack on, the metaphysical directness-claim.

The key point to be drawn from the case of the disappearing star is not to present a complete defence of this special case of the time-gap argument[82], but to show that when the focus is on a particular instantiation of the argument from causation schema, and its target is clearly stated as the metaphysical directness-claim, the argument cannot be dismissed lightly on the basis of type-2 objections as irrelevant or confused.

2.3 Conclusions

1. We have identified three types of objections to the argument from causation: type-1, that the argument is self-refuting since the conclusions undermine the premises; type-2, that the considerations adduced by the argument, with respect to the physical and physiological

---

[82] Such a defence would be complicated on two counts. Firstly, although the case of the disappearing star is generally presented on the assumption that the notion of "now" is unproblematical, the complications introduced by Special Relativity's rejection of simultaneity need to be addressed. Secondly, in common with the argument from hallucination, since the argument attempts to draw a general conclusion from a particular and unusual case some form of spreading step is required (Robinson, 1994, p. 80).

mechanisms of perception, are irrelevant to the standing of direct realism; and type-3, that

the argument depends on the Weak Phenomenal Principle.

2. Although type-1objections do present serious difficulties for the traditional form of the
   argument which takes the standard causal picture as a premise, they cannot be applied to our
   conjunctive version of the argument which claims only that the metaphysical directness-
   claim and the standard causal picture cannot both be true.

3. When, as in much of the recent literature, the argument from causation is dismissed lightly, it
   is generally on the basis of type-2 objections. However, although telling against some
   versions of the argument, these irrelevance objections are avoided when the argument is
   stated explicitly and its target is disambiguated as the metaphysical directness-claim *factually*
   construed. Moreover, such objections have no point of contact with the argument formalized
   in Section 1.3.3. Nevertheless, the strength of the case against metaphysical direct realism
   goes unrecognized when crucial distinctions are not addressed and good and bad arguments
   are rejected alike.

4. Type-3 objections are rarely found in the literature. For most direct realists, objections to the
   Weak Phenomenal Principle would be incompatible with their understanding of direct
   realism's metaphysical directness-claim. However such objections could be raised by those
   thoroughgoing reductive representationalists who are prepared to follow Armstrong all the
   way in denying that there is, in the veridical case, something of which the subject is aware.
   Here, though controversial, would be a clear-cut objection to the assertion of the Weak
   Phenomenal Principle at Premise 2 of our formalised argument. Moreover, we recognize that

it is possible, whilst accepting the Weak Phenomenal Principle, to question our call on the notion of an "immediate object of awareness" at Premises 6 and 7.

5. Our defence of the argument from causation could be challenged by those who might press the overarching irrelevance-objection, explicit in Price and Ryle but rare in the recent literature, on the basis of a demarcation between the philosophy and science of perception. From this perspective, it could be maintained that the "direct realism" of philosophy is, unashamedly, a *strictly-conceptual* thesis which stands apart from the *factual* theses properly discussed within cognitive science. We agree that "direct realism" could be defined in such terms, but maintain that such a deflationary construal would not do justice to the thesis which has been debated since the Early Moderns and is now defended within contemporary analytic philosophy.

6. Far from being the poor relation within the family of four traditional arguments, which can be dismissed lightly as irrelevant or confused, the argument from causation presents metaphysical direct realism with a serious challenge. Moreover, assuming the standard causal picture, this challenge can only be met at the expense of rejecting the notion of an immediate object of awareness or insisting on a strict demarcation between the "philosophy" and "science" of perception. For many direct realists neither option would be consistent with the thesis they are defending. Nonetheless, the two objections (as in 4 and 5 above) remain to be addressed.

Chapter 3: The debate within cognitive science

In Chapters 1 and 2 we have examined the argument from causation as it has been discussed in the philosophical literature: that is an argument which came to the fore in the seventeenth century with the picture of the physical world in which "perception depends on causal chains which link the perceiver and the object they perceive." (Robinson, 1994, p. 84) This, explicated in Section 1.3.4, we have dubbed the "standard causal picture". We maintain that the argument from causation, presented as an argument against the conjunction of the standard causal picture and the metaphysical-directness claim is valid and its premises are robust. However, notwithstanding our commitment thereto, we recognize that the premises are not entirely uncontroversial. Moreover, the controversies, with respect to intentionality and the notion of an immediate object of awareness, are substantial[83] and cannot be resolved in this dissertation. We have reached an impasse.

 To break the impasse we need to understand how the standard causal picture has been fleshed-out within cognitive science to explain the efficacy of perceptually guided behaviour and, crucially, whether or not such explanations are compatible with the claims of metaphysical direct realism. As McDermid suggests in response to Searle's dismissal of a "rather weak" causal argument, there may be versions to show that "direct realism cannot be taken seriously once we have a detailed understanding of the specific *ways* in which our perception of the physical world is causally mediated by neurophysiological processes." (2004, p. 1) In Chapter 4 we develop such an argument which shows that mainstream representation-friendly cognitive science is incompatible with the metaphysical directness claim.

Chapter 3 provides the background to this argument. It has three main objectives: firstly, to introduce the debate between the proponents of indirect and direct perception within cognitive science; secondly, to establish the connection between this debate and the contra-parallel debate within contemporary analytic philosophy; and lastly, to consolidate the alliance between the philosophy and science of perception.

In Chapter 1 we contrasted the ascendancy of direct realism within contemporary analytic philosophy, with the orthodox view within cognitive science which rejects the notion of direct perception. However it would be wrong to suggest that there is a clear-cut division on this issue.

---

[83] It is significant that Robinson opens his detailed analysis of the arguments with a discussion of the "concept of intentionality" (1994, p. 2) and that this concept plays the central role throughout. Similarly in Searle (2015), though he takes the contrary position.

As Noë and Thompson pointed out at the turn of the century, although an "orthodox " or "Establishment" view of perception "has taken shape over the last fifty years," there is a significant and increasingly influential "heterodoxy" (2002, p. 2). Importantly for our purpose, elements of this heterodoxy are sympathetic to theories of direct perception.

It would clearly be unrealistic to attempt an examination of the wide ranging discussion of perception within cognitive science, which draws on an enormous volume of work in psychology, neural science and computing. However, it is nevertheless possible to take an overview of this discussion as it relates to direct and indirect theories of perception, and then to assess the implications of these considerations for the debate within contemporary analytic philosophy, that is in terms of direct and indirect realism. Such is the objective of this chapter.

Section 3.1 introduces the debate between orthodox cognitive science and Gibson's avowedly revolutionary theory of direct perception (1968) (1979). Bruce et al (2003, pp. 301-314) provide a good introduction to Gibson's theory in their substantial overview of the "physiology, psychology and ecology" of visual perception. We focus on four points which underpin orthodox cognitive science's rejection of direct perception, and Gibson's responses thereto. Section 3.2 reviews recent developments in cognitive science, particularly with respect embodiment, computation and representation, some of which might be seen as lending support to Gibson's more controversial ideas as they have been taken up by his intellectual successors. Section 3.3 examines the connection between the debate within cognitive science and that between indirect and direct realism in contemporary analytic philosophy; representation is seen to be the pivotal issue. The section concludes by proposing a way forward in Chapter 4, which calls on a consolidation of the twenty first century alliance between the philosophy and science of perception to address the issues which remained unresolved at the end of Chapter 2. Section 3.4 summarizes our conclusions.

3.1 Helmholtzian orthodoxy v Gibsonian heterodoxy

Noë's entry on "Direct Perception" in the *MacMillan Encyclopedia of Cognitive Science* (2002) provides an overview of the debate. He opens:

> The claim that perception is direct has been advanced by a small but distinguished minority of philosophers[84] and psychologists[85] … They [supporters of direct perception] accept that perception is a form of direct access to the world, and, therefore, that the world is very much the way it seems to us in perception. For this reason, defenders of direct perception are sometimes known as naïve or direct *realists.*
>
> In fact there is nothing naïve about the theory of direct perception. Although it attempts to defend something like 'the standpoint of common sense', it is in fact a sophisticated response to the widely held view that perception could not be, in the relevant sense, direct. (Noë, 2002, p. 1)

Here Noë makes two points concerning the standing of direct perception within cognitive science. Firstly, that "direct perception", which he associates with the "'common sense'" view, is rejected by the majority. And, secondly, that this rejection of direct perception amounts to a rejection of direct realism, that is the theory espoused by "naïve or direct *realists*"; unfortunately Noë does not distinguish between metaphysical and epistemological direct realists. The first point is uncontroversial: direct perception is still the minority view within cognitive science. However, although the second point is well supported it is, as we saw in Chapter 2, disputed by those who claim that facts concerning the process of perception are somehow irrelevant to the standing of "direct realism". This issue is addressed in Section 3.3. In 3.1 and 3.2 we are primarily concerned with the debate within cognitive science, rather than its implications for the "direct realism" of analytic philosophy.

Noë and Thompson give a succinct outline of the "orthodox view" of visual perception which serves to identify four points of opposition to the notion of direct perception:

---

[84] Noë appears to underplay the support given to direct realism within contemporary analytic philosophy. See Section 1.1.

[85] He goes on "including J. L. Austin, J. Gibson, J. Mc Dowell, Ulrich Neisser, Hilary Putnam, Peter Strawson , and perhaps also Aristotle, William James and Thomas Reid."

> According to this orthodoxy, perception is a **process** [emphasis added] whereby the brain, or a functionally dedicated subsystem of the brain, builds up **representations** of relevant features of the environment on the basis of information encoded by the sensory receptors. As David Marr (chapter 11) surmises: "Vision is the *process* of discovering from images what is present in the world, and where it is." Because the patterns on the retina are **not sufficie**nt by themselves to determine the layout of the surrounding environment, perception must be thought of as a process of inductive **inference**. Perceptions are, as Richard L Gregory (chapter 7) suggests, hypotheses concerning the distal causes of proximal stimulation. In the famous phrase of Helmholtz, perception is unconscious inference. (Noe & Thompson, 2002, p. 2)

On this understanding, that is the four points of opposition and the approaches exemplified by Marr (1982), Gregory (2002) and Helmholtz (2004) it is hardly surprising that orthodox cognitive science should reject the notion of direct perception. From this perspective, the claim that perception is, in any significant sense, direct appears to be completely indefensible. Gibson, the champion of direct perception within cognitive science, would agree that, from the perspective of the orthodox view, the notion of direct perception is indeed indefensible.

Gibson is notable for his longstanding interest in visual perception, from the late 1920s through his World War II interest in depth-perception as it relates to a pilot's ability to land an airplane, and nearly three decades in the psychology department at Cornell. Moreover, he stands out within cognitive science as the preeminent advocate of direct perception. His ideas were developed over almost fifty years and presented in three major works: *The Perception of the Visual World* (1950), *The Senses Considered as Perceptual Systems* (1968) and *The Ecological Approach to Visual Perception* (1979).

It is important to note from the outset that Gibson's defence of direct perception is avowedly revolutionary and involves a radical departure from the orthodox view. He rejects not only the theories of perception posited by orthodox cognitive science, but also the construal of the problem that such theories are intended to address. Indeed, if Hermann von Helmholtz (1821 – 1894) is seen as the founding father of orthodox cognitive science, James J. Gibson (1904 – 1978) is the anti-establishment iconoclast who argued that much of what had gone before was

not only in error, but was on the wrong track and of little value (Hunt, 1994, p. 466). As Noë

puts it:

> Gibson's (1966, 1979) defence of direct perception takes as its start the rejection of the way traditional theorists frame their basic problem ….. Gibson's bold claim is that if we reformulate our analysis of the basic visual predicament [his] way, the puzzling character of how we see disappears. (2002, p. 3)

Gibson's insistence that perception must be understood in terms of its role in enabling an animal

to function appropriately (survive and thrive) in its environment, rather than a disembodied brain

discovering facts about the world, has been highly influential and widely accepted. However his

theory of direct perception was on publication, and remains, highly controversial.

3.1.1 Gibson's defence of direct perception

Drawing on Noë and Thompson (2002, p. 2), as above, we can list four points at which the

Helmholtzian orthodox view appears to be at odds with Gibson's claim that perception is direct.

1. **Insufficiency** - the information from the senses is insufficient (to account for the efficacy of perception).
2. **Inference** - inferences are involved.
3. **Process**- a complex causal chain of events is involved.
4. **Representation** - representations are involved.

Taking each point in turn we will now contrast the Helmholtzian and Gibsonian positions.

Insufficiency

According to the orthodox view, the information from the senses is "impoverished" and

insufficient to explain the efficacy of perception. As Noë and Thompson put it, "Given this

impoverished basis, how do we manage to enjoy such richly detailed visual experiences of the

environment." (2002, p. 2)

The problem is illustrated in figure 1. (Rock, 2004, p. 202) Here is an example of a situation in which, according to the orthodox view, "the patterns on the retina are not sufficient by themselves to determine the layout of the surrounding environment." (Noë & Thompson, 2002, p. 2) All three objects in the figure (imagine they are irregularly shaped tiles) give rise to the same pattern on the retina. More generally, surfaces of entirely different size, shape and orientation can project identical patterns. It is in this sense that the information in the pattern of stimulation on the retina is limited. Additional information and/or processing would appear to be required.
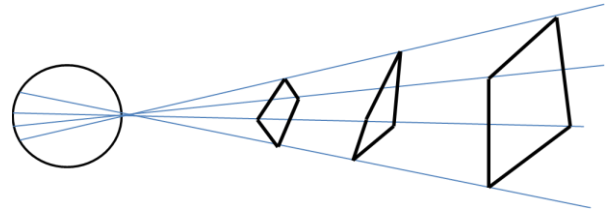
Figure 1

Gibson's response is at two levels: the first straightforward and widely accepted, the second more controversial. Firstly, he emphasizes that perception is not the processing of a stationary image on the retina, but rather an activity in which an embodied perceiver explores its environment by means of its perceptual systems. (Gibson, 1968) Such exploration involves, typically, movements of the eyes, head and body; thus much more information is available to determine the shape, size and orientation of objects. Suppose the observer were to move, then the three objects would no longer give rise to the same pattern on the retina.

However, Gibson goes much further from the orthodox view than simply pointing out that there is more information available in the flow of retinal images than had been supposed. Secondly, and from the perspective of his overall approach more significantly, he denies that the retinal image is the proper starting point for an understanding of vision. Rather, as discussed further below, the inputs to vision are "invariants" in the structured light, that is in the "optic array", and it is these invariants that unambiguously specify their sources. (1968, p. 186)

Inference

According to the orthodox view, as Helmholtz argued in his seminal work, perception involves unconscious inferences. (von Helmholtz, 2004) (Patton, 2014) That inferences are involved, follows from the first claim: if the information in the pattern of stimulation on the retina is insufficient to explain the richness and efficacy of our perceptions, it needs to be combined with information from memory in a process of inference. Moreover, according to the orthodox view,

even if the stimulus did unambiguously determine what distal objects produced it, inferential processes would still be needed to extract this information from the stimulus. As Dretske points out in this context, even if the fingerprints unambiguously identify the burglar, considerable detective work would still be required. (2004, p. 280)

However, Gibson rejects outright the claim that perception involves inference. Rather, he proposes a theory in terms of "optic arrays", "stimulus invariants", "affordances" and

"information pick up". These notions are difficult and controversial. As he admits, the laws of physical optics need to be supplemented with "the quite unfamiliar laws of ecological optics." (1968, p. 186) We can use Gibson's own diagram (p. 196) reproduce in figure 2 to get a feeling for what is involved in his account of direct non-inferential perception.
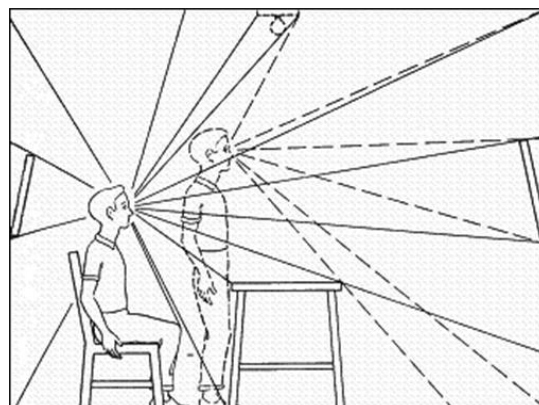
Figure 2

Consider first only the seated person and the solid lines. The solid lines to his front "represent the sample of the total optic array that is admitted to a human eye in a given posture." And the solid lines to his rear represent "the remainder of the array, which is available for stimulation but not effective at this moment." (1968, p. 195) Similar diagrams are often used in science text books to show the rays converging on the eye (geometrical optics), but here the lines do not represent light rays but rather they correspond "to the edges and corners of surfaces facing in different directions." (p. 193) This is an explanation in terms of Gibson's novel "ecological optics" which, significantly[86], he characterizes as a return to the optics of Euclid and Ptolemy. (1974)

Here we have the key notion of the optic array at a fixed point in space, from which information can, Gibson claims, be "picked up" without the involvement of any inferential processes. However he does not claim that the stationary observer has access to all the information necessary to determine the layout of the surrounding environment. Crucially important is that the observer be active and free to explore the environment. Consider now the changing optic array as

---

[86] As we noted in Section 1.3.4, Euclid proposes an extramission theory of perception.

the person stands up. Further information is now provided by those invariants, which do not change when the optic array is transformed as the observer moves within the environment. Gibson describes invariants as "non change that persists during change". (1968, p. 202) "Texture gradients", much discussed in his earlier work, are an example of such invariants. (Gibson, 1950)

The notion of "affordances", which is developed in his final work (Gibson, 1979), is a still more radical departure from the orthodox view.[87] Affordances can be seen as sophisticated versions of stimulus invariants, "invariant combinations of invariants" (p. 140), which allow higher order information about the environment to be picked up. Not only is information about physical features, say, the shape and orientation of surfaces, picked up from invariants in the optic array, but also opportunities for action which the environment affords a particular animal, such as providing a good hiding place (p. 136) or being good to eat (p. 140). Such opportunities for action are what Gibson calls affordances. Significantly, as Gibson recognizes, the key issue is not the existence of affordances but "whether [as he affirms] information is available in the ambient light for perceiving them." (p. 140) The latter claim is controversial even amongst those who are sympathetic to the notion of affordances.

Process

From the perspective of orthodox cognitive science, the most influential objection to the claim that perception is direct arises from the physical and physiological mechanism of perception. As Noë puts it in outlining, though not necessarily endorsing, the situation according to the orthodox view:

> That perception is in this way *in*direct appears to gain support from basic facts about the physics of perception. When you see a tomato you do not make direct contact with it. At best you make contact with the tomato only as mediated by a complicated causal process – the tomato affects the light which gives rise to a pattern of stimulation of the receptors in the eyes which in turn produces activity in the optic nerve and brain. At the terminus of this process there is the visual experience as of a tomato. The tomato, it should be clear, enters the process only as a more or less remote cause of the experience one eventually undergoes. (Noë, 2002, p. 2)

Here is an objection to the notion of direct perception on the basis of the standard causal picture. Indeed it is, loosely stated, a version of the argument from causation which we will examine further in Section 3.3. However Gibson rejects the standard causal picture and proposes, rather, that the perceptual systems "resonate to external information":

---

[87] As Bruce et al remark, "Gibson's approach to the psychology of perception became progressively more radical." (2003, p. 309)

> If this formula [the theory presented in the monograph] is correct, the input of the sensory nerves is not the basis of perception as we have been taught for centuries, but only the basis for passive sense impressions. These are not the data of perception, not the raw material out of which perception is fashioned by the brain. The active senses cannot simply be the initiators of *signals* in nerve fibers or *messages* to the brain; instead they are analogous to tentacles or feelers. And the function of the brain when looped with its perceptual organs is not to decode signals, nor to interpret messages, nor to accept images. These old analogies no longer apply. (Gibson, 1968, p. 5)

Gibson contrasts his own "information-based" theory of perception with existing theories within which "It is taken for granted that sensation is entailed in perception." He goes on:

> The theory of information-based perception begins with the assumption that sensory impressions are occasional and incidental symptoms of perception, that they are not entailed in perception. It is therefore not obliged to postulate any kind of operation on the data of sense, neither a mental operation on units of consciousness nor a central nervous operation on the signals in nerves. Perception is taken to be a process of information pickup. (1967, p. 162)

Here again the weight of Gibson's claim falls on his notion of "information pickup", which does all the work of the complex perceptual processes posited by orthodox cognitive science. Unfortunately, he had little to say about, and reportedly had little interest in, the mechanisms involved.[88] Nonetheless, if as Gibson supposes, "the input of the sensory nerves is not the basis of perception as we have been taught for centuries" and "the senses are analogous to tentacles or feelers", the standard causal picture is to be rejected, and he has thereby removed the major obstacle to the claim that perception is direct.

Representation

What representations are, and what aspects of the world they represent, is much debated within cognitive science. However the assumption that cognition in general, and perception in particular, involves internal representations which in some sense "stand in for" features of the external world, is essential to the orthodox view. Indeed, as we will see, when cognitive science departs from the orthodoxy of the early 1980s, this assumption is the last to be rejected. Moreover, within cognitive science, the involvement of internal representations is taken to be a hallmark of a theory of indirect perception.

---

[88] See footnote 91.

However Gibson rejects the notion of representations outright.[89] Within his account of perception, in terms of optic arrays, stimulus invariants and information pickup, there is simply no gap for representations to fill. Nakayama puts the situation well. In concluding a paper in which he gives a very positive account of Gibson's work and the developments that it has spawned, he notes:

> …I have avoided one aspect of Gibson's views that have received the greatest criticism. Gibson and especially his followers have scrupulously avoided reference to any form of internal representation. Whether this reflects a defensible ideological position as articulated by his followers (Turvey, Shaw, Reed & Mace 1981), a pragmatic ordering of research priorities as indicated by Gibson himself (1950), or a fundamental naïvete as suggested by Marr (1980, p. 30), this almost blatant disinterest in the face of steady and often brilliant progress in the field of neuroscience and psychophysics strikes me as a major limitation, particularly now. (1994, p. 329)

Thus, though highly controversial, Gibson's rejection of representations disposes of this the fourth of orthodoxy's objections to the claim that perception is direct.

At this point we can draw two conclusions from the debate within cognitive science. Firstly, from the perspective of the orthodox view the notion of direct perception is indefensible; we defer a discussion of the implications for the direct realism of analytic philosophy until section 3.3. Secondly, the claim that perception is direct can be defended, but such a defence requires a major departure from orthodox cognitive science and the standard causal picture. Just how far Gibson departs from orthodoxy can be seen in the opening remarks of a lecture he gave in 1974:

> I'm writing a book at the moment ….called The Ecological Approach to Perception …. There are three main features of this approach that I will emphasize. ….One a **new** [emphasis added] description of the environment …. to be perceived. …. Second a **new** optics which I call ecological optics to distinguish it from physical optics and geometrical optics and physiological optics which you get courses about. There are no courses in ecological optics as yet. And three, a **new** definition of perceiving. (Gibson, 1974)

3.1.2 The orthodox response

Since Gibson's ideas were a direct challenge to the Helmholtzian orthodoxy with regard to both its theories of indirect perception and its (orthodoxy's) construal of the problem to be addressed, the initial reaction to the theory of direct visual perception presented in *The Ecological*

---

[89] That Gibson did reject the notion of representation as it was understood at the time of his work is uncontroversial, and the rejection of representations is usually taken to be a defining feature of Gibsonian cognitive science. However some have interpreted him intending only to reject a certain type of representation, that is "objective" rather than "action-oriented" (Clark, 1997, p. 50) or "encoded" rather than "active" (Bickhard & Richie, 1983).

Approach to Visual Perception (Gibson, 1979) was, unsurprisingly, largely negative. Hunt

gives a flavor of the controversy that Gibson aroused:

> Such a one [a "maverick"] was the late James J. Gibson (1904-1980), whose admirers consider him "the most important student of visual perception of the twentieth century" and "the most original theoretician in the world in the psychology of perception," but whose theory is considered by the majority of perception specialists "extremely implausible" (one reviewer even called it too "silly" to merit discussion) and has few advocates. (Hunt, 1994, p. 466)

Moreover, Michaels and Carello (1981, p. 2), longstanding proponents of Gibson's approach[90],

recognize that his theory of direct perception was opposing a view of perception that "has

enjoyed nearly unanimous support". Wagemans' paper *Direct Theory of Perception: An*

Evaluation by Representatives of Indirect Theories of Perception (1986), written several years

after Gibson's last publications, provides a well referenced discussion of the response from

proponents of orthodox cognitive science. Drawing on, but not limited to, Wagemans, we can

identify four types of objection.

Firstly, Gibson cannot explain the mechanism by which direct perception is achieved.[91] Fodor

and Plyshyn put this point succinctly:

> Gibson sometimes speaks of the perceptual mechanism as "resonating" to the values of ecological parameters that they are "tuned" to. But since a more detailed account does not appear to be forthcoming, the resonance metaphor amounts to little more than whistling in the dark. (Fodor & Pylyshyn, 2002)

For Marr, who did seminal work in fleshing out a theory of perception in computational terms,

the problem is that Gibson fails to appreciate the complexity of the task:

> … although some aspects of his thinking were on the right lines, he did not understand properly what information processing was, which led him to seriously underestimate the

---

[90] "The International Society for Ecological Psychology (ISEP)" was founded in 1981 shortly after Gibson's death and is still active. Both Michaels and Carello have had a longstanding association with the society.

[91] There are stories, apocryphal perhaps, of Gibson in his later years turning down his hearing aid when questions of mechanism were raised. He certainly appeared to have little interest in the topic.

complexity of the information-processing problems involved in vision and the consequent subtlety that is necessary in approaching them. (Marr, 1982, p. 29)

Secondly, against the claim that perception is direct and unmediated, the orthodox view maintains that the perceptual process is divisible into a number of stages with definable inputs and outputs. As Ulman puts it:

…if [as he maintains exampling the case of stereoscopic vision] the extraction of visual information *can* be expounded in terms of psychologically meaningful processes and structures, then it cannot be considered immediate. (Ullman, 1980, p. 7)

Thirdly, against Gibson's aversion to representations, the orthodox view maintains that without involving the notion of representation it is not possible to account for perceptual experience or perceptually guided behaviour. Under the heading "Representation rather than resonance", Wagemans provides an extensive list of references to those who insist on the necessity of the former to account for the "experiential element in perception." With reference to perceptually guided behaviour he draws on Fodor and Pylyshyn's example[92] of navigation by means of the Pole Star:

One has to see or recognize the star as the Pole Star, to base one's behaviour on that perception. The cognitive consequences of perception are thus not only dependent on *whether* the world is seen or not, but also *how* the world is seen or represented. To explain this difference, indirect theories are based on the notion of mental representation. (Wagemans, 1986, p. 265)

Lastly, against the sufficiency and utility of stimulus invariants, the orthodox view is that the claim is unsupported by empirical evidence and that, to the contrary, the information therein (in the stimulus invariants) needs to be supplemented by information from memory and inferences thereon. As Goldstein puts it in his review of *The Ecological Approach to Visual Perception*:

The problem is that, despite Gibson's authority and the intuitive appeal of the idea of invariance, some experimental evidence must be presented[93] to support the idea that invariants are, in fact, *used* by the perceptual system. (Goldstein, 1981, p. 194)

---

[92] Fodor and Pylyshyn (2002, p. 213) discuss this issue under the heading of "The Problem of Intentionality".

[93] Gibson does discuss the "experimental evidence for direct perception" (1979, pp. 147 -187) and focuses on investigations taken to demonstrate that the observer does make use of stimulus invariants and to maintain their adequacy. However he accepts that such evidence is far from conclusive. (p. 3)

Marr makes the point in discussing texture gradients, which are perhaps the paradigm case of a

stimulus invariant:

> In an ideal world, where the surfaces are smooth and regularly and clearly marked and exhibit sufficient density of detail so that the gradients in an image can be measured quite precisely, Gibson's claim would have much to recommend it. Unfortunately, however the world is a much rougher place, in which uniformity and regularity are the exception or only an approximation rather than the rule… (Marr, 1982, p. 233)

Moreover the claim that higher order invariants, that is affordances, such as eat-ability or sit-

ability with respect to a particular organism, are specified in the optic array is seen as highly

implausible. Goldstein's response is typical:

> Gibson correctly anticipates an objection that most readers would have when he states that the skeptic may not be convinced that what food affords, something that tastes good, is specified in light. 'The taste of a thing, (the skeptic) will say, is not specified in light; you can see its form or color and texture, but not its palatability; you have to taste it for that' [Gibson, 1979, 140]. Unfortunately, he does not offer a way out for the skeptic.

> The problem comes with Gibson's statement that what an object affords is specified in the light, and his failure to deal adequately with the fact that affordances must be learned. A wooden chair may afford sitting for a human, but something to gnaw for a beaver, even though the information provided by the light is the same for both. (Goldstein, 1981, p. 193)

3.2 Recent developments

Since the publication of Gibson's last major work (1979) and the initial response in the early

1980s, there have been major advances in cognitive science some of which are relevant to the

ongoing debate between proponents of direct and indirect theories of perception. Moreover, as

argued in Section 3.3, the claim that a significant and substantial thesis of "direct realism" can

defended without any reference to these developments is highly implausible.   Here we review

these developments.

Three decades ago, Wagemans noted, with respect to Gibson's work, that orthodox cognitive

science tended to "appreciate the ecological study of the information available, without agreeing

with the theory of direct information pickup." (1986, p. 262) The same could be said today.

Gibson's ecological approach has been widely influential, whereas his theory of direct

perception remains highly controversial. However, as we will see, Gibson's intellectual successors have claimed support for the notion of direct perception on the basis of twenty first century cognitive science. Although we do not take sides in the debate, it does seem fair to say that some of Gibson's more controversial ideas, which seemed highly implausible in the early 1980s, might appear less so in the light of some of the subsequent developments.

In the following subsections we review recent developments in cognitive science under three headings: embodiment, computation and representation. The term "embodied" is sometimes used in a broad sense and contrasted with "orthodox" or "standard" cognitive science, as in Shapiro's *Embodied Cognition* (2011). However we use the term in the more restricted sense of the Gibsonian notion of an embodied perceiver in its particular environmental niche. On this understanding, we can identify two further directions of departure from early 1980s orthodoxy: the shift from the rule-based symbol-systems of GOFAI, towards connectionism and dynamical systems, and the shift from objective representations to action-oriented representations through to representation-averse theories. These directions of travel are related but need to be distinguished. Although embodiment and connectionism are often taken together, the two notions are distinct.[94]

At this point we need to refine our terminology in recognition of the fact that, although cognitive science has moved on from the early 1980s, only a small minority of cognitive scientists would defend the notion of direct perception. We will now use three terms: "orthodox" to refer to the early 1980s orthodoxy of GOFAI, embrained cognizers, and objective representations;

---

[94] It could be maintained that, whilst perception is embrained rather than embodied, the brain is a connectionist computer (Churchland, 2002). Moreover, connectionism has been associated with both representation-averse theories (Dreyfus, 2005) and representation-friendly (Churchland, 2002) theories of perception. Similarly with respect to dynamical systems, there are representation-friendly and averse interpretations. (van Gelder, 2001) Indeed the notion of representations, objective (Marr, 1982) or "action-oriented" (Clark, 1997), is retained in all but the most radical of the departures from orthodoxy.

"mainstream" to refer to the majority within twenty first century cognitive science who question some or most of 1980s orthodoxy; and "Gibsonian" to refer to the minority, notably proponents of radical embodied cognitive science (Chemero, 2009), who follow Gibson all the way in rejecting representation and espousing a theory of direct perception. However it would be wrong to suggest that they are three well defined positions, rather as discussed below, a spectrum of approaches merging one with the other.

3.2 .1 Embodiment

It is with respect to embodiment that Gibson has been least controversial and most influential. Over the last three decades cognitive science has undoubtedly shifted in Gibson's direction in recognizing that key aspects of the relationship between the animal and its environment had been seriously neglected, and that perception must be understood in the context of an embodied perceiver. However, the implications of embodiment are variously understood. As Shapiro notes, questions arise as to "whether, and if so how, the assorted views on embodiment … mark a real departure from standard cognitive science." (2011, p. 51). Similar questions arise as to whether, and if so how, the various views on "embodiment" impact on the debate between the proponents of direct and indirect perception. Embodiment cuts both ways. Indeed both directions of travel are suggested by Varela, Thompson and Rosch in their seminal monograph *The Embodied Mind*. (1991)[95]

We noted in Section 3.1.1 that the claim that perception is indirect appears to gain support from basic facts about the physics of perception which suggest that the contents of our perceptual

---

[95] That both directions might be suggested is not surprising noting the diversity of in backgrounds of the authors (Varela a biologist and neuroscientist, Thomson a philosopher, and Rosch a cognitive psychologist) and range of material introduced. Shapiro provides a useful outline. "Within embodied cognition, Varela Thompson and Rosch's *The Embodied Mind (1991)* is often regarded as an urtext. Many of its themes– disillusionment with cognitive science's allegiance to computationalism, sympathy for the Gibsonian and connectionist approaches to cognition, suspicion that the concept of representation has been overplayed in explanations of cognition, confidence that perception and cognition must be linked to action – have ascended to the status of dogma within embodied cognition. Other themes, such as the rejection of a "pre-given world", and the value of incorporating Buddhist doctrine into a science of mind have failed to garner much of a following." (Shapiro, 2011, p. 52)

experiences are dependent on the constitution of our sensory apparatus. The notion of

embodiment could be seen as re-enforcing this line of argument, by emphasizing that any

organism's perceptual capacities and experiences, including our own, are determined in large

part by the constitution of a particular species' sensory apparatus which has evolved to survive

and thrive in its particular environmental niche. As Varela et al put it*:*

> By using the term *embodied* we mean to highlight two points: first that cognition depends on various types of experience that come from having a body with various sensorymotor capacities, and second, that these individual sensorymotor capacities are themselves embedded in a more encompassing biological, psychological and cultural context. (1991, p. 173)

Chemero (1998) makes a similar point in his review of Clark's influential *Being There: Putting*

Brain, Body and World Together Again. (1997):

> [Clark suggests] that we should expect creatures (including humans) to be sensitive *only* to those aspects of their environments that matter to the actions they regularly undertake; their representations, to use Clark's phrasing, will be action-oriented. The world represented by animals with much different needs than humans will be much different than the world humans represent. This is the case because throughout their evolutionary histories, animals develop perceptual systems responsive to opportunities to fulfill their needs. Because the needs of one type of animal can be are (sic) so different from those of another, the perceptual systems that result will constitute the world in very different ways … (Chemero, 1998, p. 5)[96]

On the other hand, if the notion of embodiment is taken far enough, the shift from embrained

cognition to embodied cognition through to extended cognition, puts the perceiver in more

intimate contact with that which is perceived and could be seen as supporting Gibson's

controversial theory of direct information pickup and his analogy of perception in terms of

"tentacles or feelers". (1968, p. 5) Chemero and Silberstein identify three, increasingly radical,

departures from the image of the embrained cognizer:

> The literature makes a distinction between embodied, situated and extended cognition in supposedly ascending order of radicalness. The first claim says roughly that mind exists in the entire body, and not just in the central nervous system. The second claim says that certain environmental or social background conditions are necessary for certain cognitive

---

[96] Chemero goes on to argue that Clarks position is incompatible with the realism-claim. This is a line of argument that neither we nor Clark would support, and which Chemero (2009, p. 185) has retracted.

functions. And the third claim holds that brain-body-world are dynamically coupled and thus mental states and cognitive functions might be viewed as extending spatiotemporally beyond the skin of the organism. (Chemero & Silberstein, 2008, p. 129)

The third claim, that is "extended cognition", when it is understood in the strong constitutive sense[97] that "the mind extends even beyond the body, into the world" (Shapiro, 2011, p. 159), or "cognitive systems themselves extend beyond the boundary of the individual organism" (Wilson & Foglia, 2011, p. 3), disposes of the image of the embrained cognizer locked away from the world. Indeed, extended cognition undermines orthodoxy's master argument against direct perception on the grounds that the perceiver and perceived are separated by a complex causal chain of events. According to extended cognition, constitutively understood, the perceiver and perceived are one. (The distinction between a strong constitutive understanding of extended cognition and a weaker less radical construal thereof is discussed in Appendix A.)

The notion of an extended cognizer, and the dissolution of the distinction between perceiver and perceived, are to be found in Varela et al. In outlining their "enactivist theory", they note:

> It questions the centrality of the notion that cognition is fundamentally representation. Behind this notion stand three fundamental assumptions. The first is that we inhabit a world with particular properties, such as length, color, movement, sound etc. The second is that we pick up or recover these properties by internally representing them. The third is that there is a separate subjective "we" who does these things. (1991, p. 9)

We find a similar espousal of extended cognition and the dissolution of the subject/object dichotomy in Gibson's intellectual successors in representation-averse radical embodied cognitive science (Chemero, 2009) and ecological psychology (Oytam & Neilson, 2007). And, from a somewhat different direction, he does not altogether reject representations; Wheeler opposes Descartes' duality of subject and object, and argues for a "Heideggerian" approach to "embodied-embedded" cognitive science. (2005, p. 285)

---

[97] We use the term "constitutive" in the sense introduced by Shapiro in discussing "three themes of embodiment." (2011, p. 4) "Constitution: The body or world plays a constitutive rather than merely causal role in cognitive processing."

At this point it is sufficient to note that the Gibsonian notion of embodiment has been highly influential in twenty first century cognitive science and has important implications, albeit in a direction depending on its interpretation, for the debate between direct and indirect perception. Here we draw no conclusions except to note the implausibility of claiming that these, and the following, developments are somehow irrelevant to "the philosophy of perception." However, in Section 3.3 (and further in Appendix A) we will explore the connection between extended cognition, the rejection of representations, the rejection of the perceiver perceived dichotomy, and the standing of the claims of metaphysical direct realism: both the directness-claim and the realism-claim.

3.2.2 Computation

The second direction of departure from orthodox cognitive science is the shift from the rule-based symbol-systems of GOFAI to neural networks and dynamical systems. It is important to note that connectionist networks[98] only became a credible alternative to the rule-based symbol-system approach to cognitive science in the mid 1980s, that is some years after the publication of Gibson's final work. Dynamical systems were still later on the scene. Consequently we might expect these late arrivals to cast new light on Gibson's proposals.

In broad terms, GOFAI aligns with indirect perception. The notion of a rule-based symbol-system instantiated in the neural hardware of the brain lends support to both the epistemological claim that perception depends on unconscious inference, and the (usually implicit) metaphysical claim that the brain-bound perceiver[99] is somehow locked away from the world. On the other hand, the connectionist networks and dynamical systems approaches give reason to question

---

[98] For our purposes we can take "connectionist networks" and "(artificial) neural networks" to be synonymous.

[99] In discussing "standard cognitive science" Shapiro refers those of its proponents who maintain a "solipsistic attitude towards cognition". (2011, p. 26)

these claims. As we will see, both approaches are less clearly inferential, and the latter might be

taken to extend the cognitive agent out into the world in a strong constitutive sense.

Ramsey's *MIT Encyclopedia of Cognitive Science* entry on "Connectionism, Philosophical

Issues" provides a brief outline of three approaches:

> GOFAI [the rule-based symbol-system approach] accounts treat the mind as a complex
> organization of interacting subsystems, each performing a specific cognitive function and
> processing information through the manipulation of discrete, quasi-linguistic symbols
> whose interactions are governed by explicitly encoded rules.
>
> Connectionist networks model cognition through the spreading activation of numerous
> simple units. [The simple units are often idealized neurons, hence the term "neural
> networks".] The processing is highly distributed throughout the entire system, and there
> are no task-specific modules, discrete symbols, or explicit rules that govern the
> operations.
>
> In … the dynamic approaches to cognition, cognitive activity is understood as a series of
> mathematical state transitions plotted along different possible trajectories. Mental
> operations are described through equations that capture the behavior of the whole system,
> rather than focusing on the logical or syntactic transformations within specific
> subsystems. (Ramsey, 2001, p. 186)

The first two approaches are widely discussed and should be easily recognizable. However the

reference to "dynamic approaches" is more difficult and the question remains as to what

constitutes a "dynamical system". Van Gelder's entry on "Dynamic Approaches to Cognition"

provides a useful outline of the main difference between "dynamical work" and "classical

cognitive science" and describes a dynamical system thus:

> Most obviously, dynamicists take cognitive agents to be dynamical systems as opposed to
> digital computers. A dynamical system for current purposes is a set of quantitative
> variables changing continually, concurrently, and interdependently over quantitative time
> in accordance with dynamical laws described by some set of equations. (van Gelder,
> 2001, p. 245)

Puzzlingly, on this definition, it appears that any system that can be described in terms of

interdependent variables counts as a dynamical system. Indeed, in a more substantial treatment,

van Gelder recognizes this point when he addresses the "trivially true" objection to the dynamical hypothesis: everything is a dynamical system, so cognitive agents must be dynamical systems at some level. Van Gelder's response is that the hypothesis makes the stronger non-trivial claim that cognition "is at its highest level a dynamical phenomenon." (1998) Thus, although at the level of the physical electronics, a von Neumann computer (uncontroversially a Turing Machine equivalent) is a dynamical system, at a higher level it is a rule-based symbol-system. On the other hand the dynamical systems discussed by van Gelder are not symbol systems at any level. We can think of connectionist networks in similar terms. At the hardware level they are dynamical systems, but there is a higher level at which they are a network of weighted-input summing-and-thresholding elements, that is a network of neurons in the case of the brain.

For our purposes, the question is whether the newcomers, connectionist and dynamical systems, differ from the rule-based symbol-systems of GOFAI in ways that might be relevant to the debate between direct and indirect perception. There are five issues: capability, hardware requirements, modularity and explicability, inference, and representation. Representation is of particular relevance to our arguments and is treated in a separate sub-section.

Capability

The most fundamental question is with respect to the capabilities of connectionist networks and dynamical systems. Loosely put, can they, in principle[100], do anything that rule based symbol systems cannot? More precisely, can they exceed the capability of a Turing Machine? Although this remains an open question, the weight of opinion in computer science is that the Church Turing thesis (Sieg, 2001) implies that, in terms of the functions that can be computed, there is an equivalence between all such systems (symbol, connectionist and dynamical) and the Turing

---

[100]Connectionist networks are inherently parallel and therefore have an advantage in terms of speed for important classes of computation, notably pattern recognition, relevant to cognition.

machine. Even Copeland (2004) (2007), a notable dissenter who argues against this implication, willingly accepts that he is taking the minority position. Indeed, when he characterizes this implication, and certain corollaries within cognitive science[101], as " misunderstandings of the Church-Turing Thesis", he notes that he is in opposition to the likes of Gregory (1987, p. 784), Dennett (1993, p. 215), and the Churchlands (1990, p. 26). Moreover, in support of Gregory et al and the equivalence claim, is the fact that works on connectionist networks and dynamical systems is usually undertaken using simulations running on rule-based symbol-systems, that is the familiar PC or mainframe.[102]

Nevertheless we need to be cautious. Although the systems may all be Turing Machine equivalents in terms of the functions they can compute, there may nonetheless be significant differences in other respects. For example, as Searle points out in the context of his Chinese Room Argument (1984, p. 37), that a system can simulate a Chinese speaker by computing the same functions (appropriately mapping outputs to inputs) does not imply that the system understands Chinese; the question of understanding remains open. We return to these "other respects" in the context of representation in Sections 3.2.3 and 4.5.

Hardware requirements

As we noted above, extended cognition constitutively understood might be taken as lending support to the claim that perception is direct. Furthermore, dynamical systems might be taken as lending support to extended cognition.

---

[101] According to Copeland: "Paul and Patricia Churchland assert that Turing's "results entail something remarkable, namely that a standard digital computer, given only the right program, a large enough memory and sufficient time, can compute *any* rule-governed input–output function. That is, it can display any systematic pattern of responses to the environment whatsoever"(1990: 26)." (Copeland, 2004, p. 10)

[102] Though as Copeland notes in a section discussing the simulation of connectionist networks on von Neumann computers: "Simulations that are possible in principle are often not possible in practice. We've been able to use computer simulations up to now in our research, but only because we've been investigating relatively simple networks." (Copeland, 1993, p. 242)

Whereas rule based symbol systems and connectionist networks require special purpose hardware, say silicon chips or neurons, a much broader class of hardware can instantiate a dynamical system. On the dynamical systems understanding of cognition, no distinction need be drawn between cognitively-capable computing hardware in the organism and cognitively-incapable non-computing hardware in the environment. All or most hardware is potentially computing hardware. At its most inclusive this (controversial) notion of computation is discussed under the banner of "pancomputationalism". (Piccinini, 2015)

There are interesting connections between the dynamical systems approach and the ideas presented in Wolfram's *A New Kind of Science*, specifically his "Principle of Computational Equivalence": that "almost all processes that are not obviously simple can be viewed as computations of equivalent sophistication." (2002, p. 716) Significantly he suggests that the principle has implications "for many issues long discussed in the field of philosophy":

> … that special components are vastly less necessary than might have been thought. For it shows that all sorts of sophisticated characteristics can emerge from the very same kinds of simple components.
>
> … all sorts of systems in nature and elsewhere will inevitably exhibit features that in the past have been considered unique to intelligence – and this has consequences for the mind-body problem… (2002, p. 1196)

From this perspective, without the need for special purpose hardware, the claim that the cognitive system literally extends beyond the skin of the organism into the environment is made less implausible. (We examine the connection between dynamical systems and extended cognition further in Appendix A.)

Explicability and modularity

In discussing the orthodox response in Section 3.1 we noted that Gibson's approach was criticized on two related counts: that it could give no explanation of the mechanism of

"information pickup", and that the notion of direct perception was at odds with the fact that, according to orthodox cognitive science, the perceptual process could be analyzed in terms of a number of sub-processes. Although these criticisms still stand, a Gibsonian could maintain that, though telling when GOFAI was the only option, they lose their force in the context of connectionist and dynamical systems approaches to cognition.

Whereas early computational theories of vision, such as that of Marr (1982), posit a perceptual process which can be divided into a number of sub-processes, and can be explained in terms of the functionality thereof, there need be no such divisibility or explicability with respect to a connectionist network or dynamical system. Indeed, Chemero (2009, p. 31) suggests that non-linear dynamical systems are inherently "nondecomposable" and that their behaviour cannot be explained in terms of the functionality of their parts. Here again there is a connection with Wolfram's ideas, specifically his notion of "Computational Irreducibility", on the basis of which he suggests that even though "the components of brains follow definite laws … their overall behaviour corresponds to an irreducible computation whose outcome can never in effect be found by reasonable laws." He shows how such irreducibility arises even in simple cellular automata. (Wolfram, 2002, p. 750)

Thus, in the context of a dynamical systems based understanding of cognition, Gibson's seemingly superficial approach and failure to provide an explanation of the mechanism of "information pickup" might be regarded as essential features, rather than shortcomings.

However, although such considerations might be taken up in defence of Gibson's holistic approach, recent developments in neuroscience give increasing support to modular explanations of the perceptual process at the physiological level in terms of the functionality of particular

identifiable regions of the brain such as V1, V2 etc. (Bruce, et al., 2003) On the basis of this neurophysiological evidence it seems that, though cognition could perhaps be instantiated in a number of connectionist networks or dynamical systems corresponding to V1, V2 etc, the perceptual process is nonetheless decomposable into, and explicable in terms of, sub-processes with identifiable inputs and outputs.

The role of inference

For a proponent of Gibsonian cognitive science, who wants to avoid the objection that perception involves inference, the attraction of the connectionist network is that it computes but does not proceed by inferential steps; it might appear to offer computation without inference. As Shapiro puts it:

> … despite the ease with which the [connectionist] network might be described as computing functions, there is no obvious need to conceive of the network as availing itself of rules of inference, of testing hypotheses, of making assumptions, and so on. The idea that that there is somewhere in the network and algorithm that dictates the manipulation of symbols also seems out of place. (2011, p. 45)

There is a sense in which rule-based symbol-systems do, and connectionist networks do not, make use of "rules of inference". If we take "rules of inference" to be instantiations of the "if then …else …." structures of GOFAI's rule based systems, they are not to be found in connectionist networks; the same point could be made with respect to dynamical systems. However, it is questionable whether this is the sense of "inference" on the basis of which orthodox cognitive science rejects the claim that perception is direct. It seems that, "rules of inference" in the sense at issue between Helmholtzian and Gibsonian cognitive science, say rules relating visual angle and distance (Rock, 2004, p. 208) (Marr, 1982, p. 156), could equally well be instantiated in algorithms in a rule-based symbol-system or connection weights in a connectionist network.

3.2.3 Representation

The notion of representation is not only central to the debate within cognitive science, it has also been at the core of the discussion of perception within Western Philosophy since the Early Moderns. Indeed Churchland draws the parallel between the term "representation" as it is used in the contemporary literature and "idea" in Hume, "thought" in Descartes and "concepts" in Kant. ""Representation" is just the term currently in fashion." (2002, p. 275) Unsurprisingly then, recent developments in our understanding of representations and the instantiation thereof, notably in brain-like systems, are significant with respect to the question as to whether, and in what sense, perception could be direct.

Of the three directions of travel from orthodoxy in the direction of the Gibsonian heterodoxy, the rejection of representations is the most controversial. Even those who espouse embodiment and dynamical systems, in the main, remain committed to the notion of representation. For most, as Ramsey puts it, "It has become almost a cliché to say that the most important explanatory posit today in cognitive research is the concept of representation." (2007, p. xi) And according to O'Brien and Opie (2009), representation is "the only good idea we have ever had about how intelligence might arise in a natural system." Representation is certainly the most widely discussed concept in the recent literature.[103]

As we have seen, the claim that the "brain ... builds up representations of relevant features of the environment" (Noë & Thompson, 2002, p. 2) is essential to both (early 1980s) orthodox and (contemporary) mainstream cognitive science. Moreover, up until the mid 1980's, such representations, instantiated as data structures within the rule-based symbol-systems of GOFAI, were generally taken to be digital rather than analogue, static rather than dynamic, local rather

---

[103] Eg (Cummins, 1996) (Dretske, 1997) (Edelman, 1998) (Clark, 1997) (Churchland, 2002) (Kulvicki, 2004) (Wheeler, 2005) (Ramsey, 2007) (Chemero, 2009) (O'Brien & Opie, 2011) (Shapiro, 2011) (Churchland, 2012). Less recent but particularly useful (Palmer, 1978)

than distributed, and objective rather than action-oriented. However, over recent decades, partly arising from work on connectionist networks and dynamical systems, and partly from the espousal of the notion of embodiment, different types of representation have been posited. Thus although the three types of computation are generally regarded as Turing Machine equivalents in terms of the functions they can compute, there may nevertheless be differences with respect to their representational capabilities.

Digital v analogue

In contrast to rule-based symbol-systems, the variables in connectionist networks (input weightings and cell activation levels) and dynamical systems (more generally, system states) are continuous (analogue), rather than discreet (digital). However, as van Gelder suggests, the treatment of time may be a more important issue: "The time variable in dynamical models[104] is not a mere discrete order, but a quantitative, sometimes continuous approximation to the real time of natural events." (1998, p. 245) Thus the system state evolves continuously rather than in discreet time steps. On both counts, dynamical systems and some connectionist networks might be taken to have a more fine grained representational capability than that offered by rule-based symbol-systems. Moreover, "continuous … real time of natural events", in contrast to "mere discreet order", has something of a Gibsonian ring; consider Gibson's notion of "optic flow". (1979, p. 223)

Less well recognized, but seemingly more important, is the distinction drawn by O'Brien and Opie. They argue that in analogue computation, but not digital, there is a "physical analogy" between the "target system" and the "representing vehicle":

> An analog computer is ''governed by'' or ''obeys'' mathematically expressible laws that are similar to those governing the represented system. This does not apply to a digital computer, whose governing equations are, in most cases, those that apply to the

---

[104] The same could be said for re-entrant connectionist networks. (Jordan, 2001, p. 709)

components of a digital electronic device. A digital computer, when being used to model some target system, does so by ''running the maths'', that is, by producing or calculating the values of a mathematical function that describes the target process or object. It does not itself obey those equations, but rather those that govern the dynamics of its representing medium. (O'Brien & Opie, 2009, p. 8)

Drawing on this distinction they maintain that in connectionist networks, but not rule-based symbol-systems, representational content is an "intrinsic property" of the "representing vehicles". (O'Brien & Opie, 2001) Thus connectionist theories (qua analogue computation) provide a better basis for a "naturalistic explanation of phenomenal consciousness". In particular, connectionist theories are able to resolve certain objections to the "representationalist theory of consciousness".[105] We will return to this point in Section 4.5.

Static v dynamic

The representations posited in orthodox cognitive science are static in the sense that they are instantiated in a data structure fixed at a particular point in time. Similarly, static representations could be instantiated in a set of cell activation levels or systems states in a feed-forward connectionist network. However the notion of dynamic representations which are instantiated in a, perhaps repetitive, sequence of states has now been introduced. According to van Gelder, an advocate of the dynamical systems approach to cognition, dynamical models usually "incorporate representations, but reconceive them as dynamical entities (e.g., system states, or trajectories shaped by attractor landscapes)." (2001, p. 245) Significantly perhaps, in the context of Gibson's call on the notion of resonance, proponents of adaptive resonance theory (ART) have used dynamical systems as perceptual models within which "a "resonance loop" is established between some sensory feature units and a particular classification unit." (Port, 2002)

---

[105] On the basis of the "*classical* computational theory of mind" the representationalist theory of consciousness cannot account for: (1) the distinction between representational content of which we are, and are not, conscious, and (2) the causal efficacy of representations - this requires that representational content must be an intrinsic property of "the brain's representing vehicles". (O'Brien & Opie, 2001)

Local v distributed

As van Gelder points out there is no sharp distinction between local and distributed representation; it is more a matter of degree. (2001, p. 236) However, in simplistic terms, the supposed "grandmother neuron" (Gross, 2002) which fires whenever grandmother comes into the visual field would be an example of local representation, whereas a collection of neurons which respond with a particular pattern of firings would be an example of distributed representation. Although not always clear-cut, the distinction can be useful. The representations in symbol-systems are often characterized as local, whereas distributed representation is taken to be a feature of the hidden layers of connectionist networks, and of dynamical systems.

Objective v action-oriented

The three extensions of the orthodox notion of representation so far discussed have arisen from the advances in computing and the recognition that GOFAI is not the only game in town. However the notion of action-oriented representations arises from the Gibsonian notion of embodiment and the emphasis on an animal exploring its environment. Moreover the distinction between objective and action-oriented representation concerns what is represented, rather than how it is represented. Chemero cites Clark (1997) as the first to coin the latter term which he introduces thus:

> Action-oriented representations differ from representations in earlier [objective] computationalist theories of mind [notably Marr (1982)] in that they represent things in a non-neutral way, as geared towards an animal's actions, as affordances. … the perceptual systems of agents need not build an action neutral representation of the world, which can then be used by the action-producing parts of the agent to guide behaviour; instead, the agent produces representations that are geared towards the actions it performs from the beginning. (Chemero, 2009, p. 26)[106]

---

[106] Similarly, Wheeler describes objective representations as "essentially objective, context-independent, action-neutral, stored descriptions of the environment." On the other hand action-oriented representations are more akin to "egocentric control structures for situation-specific actions". (2005, p. 196)

The key point at this stage in our discussion is that the orthodox notion of representations as digital, static, local and objective has been greatly extended over the last three decades. We will return to these developments in our detailed examination of representations in symbol-systems and connectionist networks in Chapter 4.

Representation-averse approaches

We now turn to developments that might lend support to representation-averse theories of perception. Although not entirely uncontentious, for example Ramsey (2007) argues to the contrary[107], we will take it that all three computing paradigms can instantiate representations, albeit there are differences therein. However the question remains as to whether connectionist networks and dynamical systems need to do so; the notion of representation would appear essential to that of a "symbol" in a symbol-system. It is possible that although connectionist networks and dynamical systems can instantiate representations, it may be the case that they can also provide the basis for a representation-averse theory of perception. Indeed, although van Gelder notes that dynamical systems are usually taken to involve representations, he allows that "some dynamicists claim to have developed wholly representation free models". (2001, p. 245)

Dreyfus, taking his cue from the Phenomenological Tradition, argues that connectionist networks might also provide representation free models of cognition. As he puts it, although Merleau-Ponty argued persuasively against models of mind or brain function that were based on representations, until recently there did not appear to be viable alternatives:

> Fortunately, however, there are now models of what might be going on in the brain of an active perceiver forming an intentional arc that do not introduce brain representations. Such models are called simulated neural networks. (Dreyfus, 2005, p. 133)

Dreyfus is not alone in drawing such a connection and we will consider the affinity between Gibsonian cognitive science, representation-averse theories and the Phenomenological Tradition

---

[107] Ramsey maintains a particularly strict criterion for what he regards as genuine representations which can only be met by the "Classical Computation Theory of Cognition". (2007, p. 1)

in Section 3.3. Significantly, the representation averse-approach has also drawn support from a quite different direction, that is robotics. Here Brooks is known for his maxim "Intelligence without representation", and for building autonomous robots which, in distinction to the standard AI approach, do not incorporate a representation of the world. According to Brooks, "It turns out to be better to use the world as its own model." (1991, p. 139) Shapiro makes clear the relevance of this work to the Gibsonian approach to perception:

> Brook's behavior-based robots might for good reason be called "Gibsonauts", for their design embodies a Gibsonian approach to perception. Sensing feeds directly into action without a representational middleman. No inferential steps from an impoverished input to a perceptual output are necessary. Gibson would have approved of Brook's design …. (Shapiro, 2011, p. 139)

We do not take sides in the debate between representation-friendly and representation-averse cognitive science. The key point at this stage is that, whereas at the time of publication Gibson's representation-averse theory of direct perception (1979) stood alone against an ascendant GOFAI, it can now draw support from, albeit a minority, within cognitive science.

As we have emphasized, the rejection of representations is the most controversial of the departures from orthodox cognitive science. Indeed, the minority who do follow Gibson in rejecting the claim that perception involves representations advertise their position as "radical embodied cognitive science", in contradistinction to mere common or garden "embodied cognitive science" which is regarded as "non-radical" by virtue of its failure to reject the notion of representation. (Chemero, 2009) (Hutto & Myin, 2012)

Radical embodied cognitive science can be understood as being at the heterodox end of a spectrum of views, with the orthodox view at the other. According to paradigmatically orthodox cognitive science, perception is a process wherein the brain builds up objective representations of the world on the basis of rule-based symbol-manipulation of information from the senses. Marr's computational theory of vision (1982) is towards this end of the spectrum and conceives

of representations as data structures, though he is not committed to the rule-based symbol-systems of GOFAI. The approach of, say, the Churchlands (2002) (2012) is mid-spectrum. They espouse brain-like connectionist networks, but retain representations, now distributed, as an essential explanatory concept. Cognition is still essentially embrained rather than embodied or extended. Further from the orthodox end, proponents of non-radical embodied cognition, such as Clark (1997) and Wheeler (2005), espouse embodiment and follow Gibson in emphasizing that perception involves the whole organism actively exploring its environment, but again retain representations, albeit they are now explicitly action-oriented rather than objective. It is only at the extreme end of the spectrum that proponents of radical embodied cognitive science, notably Chemero (2009), follow Gibson all the way in championing direct perception and rejecting the notion of representation outright.

3.3 The connection with the direct realism of analytic philosophy

We have seen that although there is an influential Gibsonian minority which might claim support from some of the developments discussed above, mainstream cognitive science remains true to its Helmholtzian roots in rejecting the claim that perception is direct. The question now arises as to the connection between the notion of direct perception rejected by mainstream cognitive science and the "direct realism" defended within analytic philosophy. Here opinions divide.

Some take cognitive science's rejection of direct perception to constitute a rejection of direct realism. For example, Noë takes the terms to be synonymous: "defenders of direct perception are sometimes known as direct or naïve *realists*." (2002, p. 1) Moreover the indirect perception espoused by mainstream cognitive science is widely interpreted as a version of indirect realism. For example Maund suggests that Rock's "version of indirect realism", as discussed below, is close to that defended by Descartes and Locke. (2003, p. 14) On this understanding, mainstream cognitive science and much of contemporary analytic philosophy are in opposition. Indeed,

Smythies charges "many contemporary philosophers" with adhering to the "prescientific theory of 'commonsense'- naïve or Direct realism." (1993, p. 205)[108]

On the other hand, as we have seen, others[109] have distanced the "philosophical thesis" of direct realism from those considerations, with respect to the mechanism of perception, which motivated indirect realists from the Early Moderns, through to the likes of Smythies, in rejecting the notion of direct perception. On this understanding cognitive science and philosophy are addressing different types of question, in our terms *factual* and *strictly-conceptual* respectively, and any conflict is only apparent.

We will argue that it is impossible to maintain a distinction between the questions proper to the "philosophy" and "science" of perception. Indeed, the recent developments with respect to embodiment, computation and representation would cut across any such divide. Peaceful coexistence, on the basis of a withdrawal into separate magisteria[110], is not possible. Thus there is a prima facie conflict between contemporary philosophy's defence of direct realism and mainstream cognitive science's rejection of the claim that perception is direct. However exactly what is at issue is less clear and needs careful examination.[111] To this end, drawing on the distinctions introduced in Chapter 1, we examine the implications of the debate within cognitive science with respect to direct realism's directness and realism claims. The cognitive science, no less so than the philosophy, is compromised by a failure to maintain these crucial distinctions; indeed the tendency to conflate "realism" with "direct realism" is widespread, as in Gibson's *New Reasons for Realism* (1967).

---

[108] He argues that "neuroscience and introspection psychology offer a complete account of perception and consciousness that leaves no room whatsoever for the theory of Direct Realism." (1993, 205)

[109] Notably (Price, 1950) (Ryle, 1953) (Hamlyn, 1961) (Locke, 1967) (Fumerton, 1985)

[110] In the sense of Gould's non-overlapping magisteria. (1999)

[111] For example, according to Dretske, "The sense (if any) in which this [Gibson's] theory is direct is much different from the sense in which Direct Realism is direct. Direct Realism is a theory about the objects of perception, about what we see. The kind of direct realism we are now talking about, the kind associated with Gibson's work, is a theory about the processes underlying perception, about how we see what we see." (Dretske, 2004, p. 281)

3.3.1 The epistemological directness-claim

As discussed in Section 3.1, from the perspective of orthodox cognitive science, the claim that perception is direct faces objections with respect to: (1) the insufficiency of the information from the senses, (2) the involvement of inferences, (3) the involvement of a complex causal chain of events (4) the involvement of representations. The question is whether, and how so, these objections, and Gibson's counters thereto, bear on direct realism's directness claims.

Objections (1) and (2) are seemingly at odds with the epistemological directness-claim, which denies that our perceptual judgments about the physical world are based on inference. Indeed, objection (2) which follows from objection (1), though stands independently[112], appears to be a clear cut objection to the epistemological directness-claim. This conflict is to be expected. Helmholtz, often taken to be the founding father of orthodox cognitive science, famously describes perception as "unconscious inference" (Patton, 2014). If perception involves inference, as claimed by mainstream cognitive science, then epistemological direct realism must be false. We appear to have reached a conclusion with very little effort. However it might be argued that the inferences espoused by mainstream Helmholtzian cognitive science are not those which epistemological direct realism rejects.

We would argue, to the contrary, that the type of inferences which mainstream Helmholtzian cognitive scientists, such as Rock, Gregory and Marr, posit in their respective theories of perception, are clearly of a type that an epistemological direct realist would need to reject.[113]

---

[112]As we noted in Section 3.1,according to the orthodox view, even if the stimulus did unambiguously determine what distal objects produced it, inferential processes would still be needed to extract this information from the stimulus

[113] For example Rock explains size constancy in terms of "the rule … that the visual angle subtended by an object is an inverse function of its distance … [from which] size perception can be inferred given the visual angle and the apparent distance of the object." (Rock, 2004, p. 208) Moreover we would argue that the rule, and inferences thereon, could equally well be instantiated in the weights and connectivity of a connectionist network, or the differential equations of a dynamical system, as in the algorithms and symbols of GOFAI.

But, since the standing of epistemological direct realism is peripheral to our central thesis, we do not need to establish this claim.

However Gibson would come to the rescue of the epistemological directness-claim. As we saw in Section 3.1.1, Gibson disarms objections (1) and (2) to the notion of direct perception. He denies that the information from the senses is impoverished, and denies that, for this or any other reason, perception involves inference. Rather, visual perception is to be understood in terms of direct and non-inferential information pickup from the optic array.

Unsurprisingly then, as discussed further below, Gibson believes that his theories lend support to epistemological direct realism:

> The availability of information in ambient light and the possibility that it can be picked up directly have implications for epistemology. They lend sophisticated support to the naïve belief that we have direct knowledge of the world around us. They support direct realism. (Gibson, 2002, p. 89)

3.3.2 The metaphysical directness-claim

It is with respect to points (3) and (4), process and representation, that mainstream cognitive science appears to oppose the metaphysical directness-claim which maintains that the immediate object of awareness is one and the same as the physical object. Indeed the orthodox view that perception is "a process whereby the brain ... builds up representations of the relevant features of the environment on the basis of information encoded by the sensory receptors" (Noe & Thompson, 2002, p. 2), can be seen as endorsing and fleshing-out the naturalistic understanding of the perceptual process which motivated the indirect realism of the Early Moderns. Moreover when perception is addressed from the perspective of contemporary cognitive science, the physics and physiology of the perceptual process come to the fore and demand attention. As Noë puts it in outlining, though not necessarily endorsing, the objections to direct perception, the claim that perception is indirect appears to gain support from the physics of perception: "The

tomato, it should be clear, enters the process only as a more or less remote cause of the experience one eventually undergoes. [quoted more fully in Section 3.1.1]" (Noë, 2002)

That the tomato is a "remote cause" of the experience of seeing the tomato, clearly suggests a prima facie objection to the thesis that the immediate object of awareness is one and the same as the physical object: if A is the remote cause of B, then A cannot be identical with B. This is, loosely stated, a version of the argument from causation against metaphysical direct realism. As Noë puts it, it is a "philosophical line of argument, but one that has been entrenched in scientific theories of perception". (2002) Moreover, as Searle notes, this line of argument has, in the twentieth century, gained in appeal from "the prestige of the natural sciences." (1998, p. 28) As we noted in Section 1.1, much of contemporary natural science appears to assume what is, in effect, a loosely formulated version of Lockean indirect realism.

Nevertheless, as we have seen, Searle, along with the majority of contemporary philosophers, rejects the argument from causation. Despite its ever greater prestige and influence in the twenty first century, a direct realist might still maintain that cognitive science has, by merely endorsing and fleshing-out the standard causal picture, added nothing to the argument from causation as it was propounded by the Early Moderns, and has done nothing to address the objections thereto. As Hamlyn puts it, the findings of cognitive science are "essentially *of the same kind* as those of the 17th century, or indeed the Presocratics." (1961, p. 185)[114] More recently, Bennett and Hacker characterize Smythies' call on the findings of neuroscience to support his objections to direct realism as a "bad piece of seventeenth century metaphysics, which is not improved by a smattering of up to date neuroscientific discoveries about perception." (2011, p. 249)[115]

---

[114] In the terminology of the distinction introduced in Section 1.2, Hamlyn takes the matters at issue to be *strictly-conceptual* rather than *factual*. He continues: "Because the problem is general, ie conceptual, no greater resort to the details of factual knowledge will provide fresh insight." (Hamlyn, 1961, p. 185) It is significant that Hamlyn cites Hirst (1959), who we quote approvingly, as taking the contrary position.

[115] See Smythies (2011, pp. 252-254) for his response.

We take issue on both points. If we understand the "seventeenth century metaphysics" to be a reference to the argument from causation, we have endorsed it (the metaphysics) in Chapters 1 and 2, and moreover in Chapter 4 we will maintain that it is improved by "up to date" cognitive science. However, our focus will be on representation and computation rather than neuroscience. Indeed representation will now take centre stage in this dissertation. (This is not to say that proponents of indirect realism could not argue their case on the basis of other developments in cognitive science. For example, Smythies and Ramachandran (1997) claim to present a conclusive refutation of the direct realist theory of perception on the basis of experimental psychology.)

In cognitive science, the involvement of representations is taken to be one of the hallmarks of an indirect theory of perception. Moreover, on the basis of the orthodox view, that the brain builds up representations of the relevant features of the environment, it is plausible to reify representations as the immediate objects of awareness. Consequently, objection 4, the involvement of representations, lends prima facie support to the argument from causation by offering representations, as a naturalistically acceptable alternative to sense data, as the terminus of the perceptual process.

With respect to the involvement of representations, direct realists divide. As we have seen in Section 2.2.3, some disjunctivist direct realists maintain that the involvement of representations is at odds with a construal of "direct realism" which does justice to the immediacy of perception. On this basis there appears to be a clear-cut disagreement between those disjunctivist direct realists who deny the involvement of representations tout court and mainstream representation-friendly cognitive science, though nothing rests on this suggestion.

However representationalist direct realists see no such problem, and have no compunction in accepting that representations play an essential role in perception. According to Dretske, "We perceive a cat by (internally) representing a cat, not by perceiving an internal representation of a cat." (2004, p. 276) Indeed Strong Representationalism accounts for the directness of perception in terms of an identity between phenomenal character and representational content. Moreover, we accept that, although robust, our formalized argument presented in Section 1.3.3 remains exposed to an objection to the assertion of the Weak Phenomenal Principle at Premise 2, which could be made by a proponent of Strong Representationalism who is prepared to deny the existence of an immediate object of awareness even in the veridical case.

However, the impasse can now be broken. An examination of the various representations discussed in detail within cognitive science provides the basis for a more sophisticated version of the argument from causation which does not draw on the Weak Phenomenal Principle and, thus, provides a robust objection to representionalist/intentionalist direct realism. The key point is that, by drawing on the cognitive science to understand what representations actually are, what they represent and how they represent, our argument does not rest on the mere involvement of representations, but rather on the relationship between representations and their targets and thus the constraints on the contents of such representations. In response to Dretske, we maintain that the constraints on the representational content provided "by (internally) representing a cat" are incompatible with the content that would be required to satisfy the claim that we directly perceive the cat, that is directly perceive in the sense of the metaphysical directness-claim.

Although we reject his representationalist direct realism, we endorse Dretske's understanding of the questions that need to be answered. In the context of discussing the debate within cognitive science and the significance of Gibson's theory of direct perception, he notes:

> It is true, but unilluminating, to be told that the sensory perception of an object occurs when the visual system constructs a sensory representation of the object. What we want to know is what kind of representation the sensory representation is. …. Is it something

like what philosophers and psychologists used to call a sensation? Or is it more like what they (or some of them) now call the percept? Or, to use even more fashionable jargon, is it more like what Marr (1982) and his associates call a 2½-D sketch? …. Until these questions are answered, we can expect little progress on questions about the nature of perception itself. (Dretske, 2004, p. 282)

And, in a paper defending his Strong Representationalist theory[116] of perception, Dretske clearly recognizes the difficulties in so doing:

One thing that might (if it were true) tell against a representational theory is if there were no plausible theory of original or natural representation that could put some flesh on these bones. …There are of course philosophical theories of representation (including my own: see Dretske 1995) that purport to give a naturalistic account of representation …. But there isn't much agreement (at least not *positive* agreement) about the plausibility of such theories. (Dretske, 2003, p. 74)

In Chapter 4 we draw on an examination of the representations posited within cognitive science to argue that there is no prospect of such a "plausible theory" which is compatible with the metaphysical directness-claim. However, although we take issue with Dretske on this point, we are sympathetic to his approach and his understanding of the issues involved.[117] Indeed the issue that stands between us is clear-cut: we maintain that none of the representations so far posited, or in prospect, can supply the representational content required of a representation in a direct realist theory of perception.

We now turn to the Gibsonian view. As we saw in Section 3.1, Gibson disarms objections to direct perception by rejecting both the understanding of the perceptual process that "we have been taught for centuries" (1968, p. 5) and the involvement of representations. Thus, although the extent to which mainstream cognitive science adds weight to the argument from causation might be disputed, Gibsonian cognitive science clearly undermines the argument by rejecting the standard causal picture on which it (the argument) is grounded. Moreover, the particular

---

[116] That is: "the claim that perceptual experiences are not only representational, but that their phenomenal character – the qualities that determine what it is like to have the experience – are *completely* given by the properties the experience represents things to have." (Dretske, 2003, p. 67)

[117] Significantly perhaps, both Dretske and the author of this dissertation studied electrical engineering before turning to philosophy.

difficulties for direct realism that have been taken, by some, to arise from the involvement of representations are avoided altogether. Thus Gibsonian cognitive science defeats the putative defeaters of the metaphysical directness-claim. (However, as we will see in Section 3.3.3, Gibson's avowedly revolutionary approach raises issues with respect to the realism-claim; here again representation will be the pivotal issue.)

As we noted in Section 3.3.1, Gibson believes that his theories lend support to direct realism. He discusses the philosophical implications of his work most fully in his paper *New Reasons for Realism* (1967):

> It seems to me that these hypotheses make reasonable the common sense position that has been called by philosophers direct or naive realism. I should like to think that there is sophisticated support for the naive belief in the world of objects and events, and for the simple-minded conviction that our senses give knowledge of it. But this support is hard to find when the senses are considered as channels of sensations; it becomes easy when they are considered as perceptual systems.
>
> Highly ingenious philosophical arguments have been advanced in this century that give roundabout support for the common man's position. It is my impression (although I could be wrong) that all these forms of realism presuppose what I have called the theory of sensation-based perception, and that this is why the arguments have to be roundabout. (Gibson, 1967, p. 168)

Here Gibson makes two points: firstly that his theory of perception supports direct realism, and secondly that on the basis of "the theory of sensation-based perception", in our terms the standard causal picture, direct realism is hard to defend. We take each in turn.

Although his references to epistemology, belief and knowledge, suggest that Gibson is concerned with epistemological rather than metaphysical direct realism, the distinction is blurred in his subsequent discussion. He explains how his theory supports direct realism under three headings: "Immediate or Direct Experience", "The Detection of Colors or Sounds" and "Public Experiences and Private Experiences". Under the first heading, after noting that the orthodox sensation-based doctrine implies that "our experience of objects and events is indirect", he continues:

> For this doctrine we now have a substitute. There can be direct or immediate awareness of objects and events when the perceptual systems resonate so as to pick up information *and* there can be a kind of direct or immediate awareness of the physiological states of our sense organs when the sensory nerves as such are excited. But the two kinds of experience should not be confused, for they are at opposite poles, objective and subjective. (Gibson, 1967, p. 168)

Here Gibson's rejection of the presupposed "theory of sensation-based perception" in favour of his own avowedly radical alternative, could be seen as allowing the direct realist to avoid conceding that perception is, in any significant sense, epistemologically or metaphysically, indirect.

Under the second heading, Gibson sketches out how his theory can be used to support the "man-in-the-street" against the indirect realist philosopher, who bewilders the "poor man" by asserting that "colors are only in *him* [the perceiver] since light consists of waves". Gibson maintains that ecological optics supports the plain man, and concludes that the "doctrine of secondary qualities comes from a misunderstanding."[118] (Gibson, 1967, p. 170) Here, in addressing the argument from secondary qualities[119], Gibson appears to be defending metaphysical direct realism. Under the third heading, Gibson again draws on ecological optics and acoustics, to maintain, against the indirect realist, that two or more perceivers can "have the same perceptions". The argument is essentially that, by moving about the environment, the perceivers can all have access to the same information in the "ambient stimulation". (p. 171) Here again Gibson could be understood in lending his support to metaphysical direct realism.[120]

---

[118] However the precise line of argument in his brief explanation is not clear, but appears to be on the following lines. The reflectances and pigmentation of surfaces determine the structure of ambient light and hence the stimulus information therein. Colour is then an invariant of structure, which specifies "classes of natural pigments". And information about the environment can be picked up from such invariants. He concludes that "In short, there is a proper meaning of the word 'color' that refers to a distinctive feature of a solid substance." Gibson argues similarly with respect to sound.

[119] The argument that begins from the premise that science has shown that physical objects do not posses secondary qualities, such as colour, intrinsically. (Robinson, 1994, p. 59)

[120] That is, by countering the possible objection that, since the immediate object of awareness is not the same for two perceivers perceiving the same physical object, no such immediate object of awareness can be one and the same as the physical object.

However, although we maintain that Gibson's theory of direct visual perception does, by virtue of rejecting the standard causal picture, lend support to the metaphysical directness-claim, it would be wrong to suggest that Gibson is explicit on the issue. Indeed Gibson rarely addresses the metaphysical implication of his avowedly radical approach to perception.[121] As Sanders puts it in a paper drawing parallels with the work of Merleau-Ponty, "In Gibson's case the metaphysical/epistemological implications were almost never explicit". (1993, p. 289) However, as we will see, Gibson's intellectual successors do engage with these issues.

We now turn to the second point. Gibson implies that direct realism is difficult or impossible to defend when, as in the theory which he takes to be held by the philosophers, perception is assumed to be "sensation-based". Indeed he characterizes such defences as "ingenious" [122] and "roundabout" (1967, p. 168), and later implies that they are unsuccessful. In contrast he suggests that his information-based theory of perception will provide a more obvious and direct defence of direct realism. The suggestion is, in our terminology, that if the philosophers were to reject the standard causal picture, direct realism would be an easier case to argue.

We find a similar point voiced by Gibson's intellectual successors. From the perspective of radical embodied cognitive science, Chemero puts it thus: without "a new ontology, one that is at odds with today's physicalist reductionist consensus … direct perception is simply indefensible". (2009, p. 136) And, from the perspective of Gibsonian "Ecological Psychology", according to Turvey, on the assumptions of local causation, which we take to be essential to the standard causal picture, "James Gibson hasn't got a hope, you cannot get an intuition about direct

---

[121] Dretske argues that Gibson's theory does not address the question of what we perceive but, rather, how we perceive. Hence the "sense (if any) in which this theory is direct is much different from the sense in which Direct Realism is direct." (Dretske, 2004, p. 280)

[122] Perhaps he had in mind the early adverbialist, representationalist and disjunctivist responses to the arguments from illusion and hallucination.

perception from this frame of reference." (Turvey, 2003) (We discuss these issues further in Appendix A.)

3.3.3 The realism-claim

Mainstream cognitive science does not question the realism-claim, indeed it is grounded on the assumption that the physical world has an existence that is not in any way dependent on it being cognized. When mainstream cognitive science takes issue with direct realism, it is the claim that perception is direct, rather than that there is a perceiver independent world, that is at issue. Although mainstream cognitive science may not be direct *realist* it is nevertheless *realist*; it has no truck with idealism. However, at the more heterodox end of Gibsonian cognitive science, there are some who explicitly reject the realism-claim (Varela, et al., 1991), or espouse a Berkeley-type "realism" (Oytam & Neilson, 2007), and others who, at least, recognize that the realism-claim is in danger of being compromised and needs to be defended albeit in a qualified form (Chemero, 2009) (Zahidi, 2013).

As we have seen Gibson proposes an ecological approach to perception which, being both direct and non-representational, can be seen as the precursor of today's radical embodied cognitive science. The anti-realist leanings of such an approach are particularly evident in Gibson's later work[123], as seen in his discussion of "affordances", a notion which plays a key role is his most fully developed theory of perception:

> … an affordance is neither an objective property nor a subjective property; or it is both if you like. An affordance cuts across the dichotomy of subject – object and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behaviour. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer. (Gibson, 1979, p. 129)

Gibson emphasizes the extent to which his thinking departs from the orthodox view:

> The theory of affordances is a radical departure from existing theories of value and meaning. It begins with a new definition of what value and meaning *are*. The perceiving

---

[123] Bruce et al outline the manner in which Gibson became "progressively more radical". (2003, p. 309)

of an affordance is not a process of perceiving a value-free physical object to which meaning is somehow added in a way that no one has been able to agree upon; it is a process of perceiving a value rich ecological object. (p. 140)

That there is a tension between Gibson's ecological approach and the reality-claim is strongly suggested by the parallels that have been drawn, by his own supporters, between Gibson and thinkers who have been taken as hostile to scientific naturalism or the subject/object distinction. For example, Sanders emphasizes the connection between Gibson and Merleau -Ponty. In a paper drawing parallels between Merleau-Ponty's "materiality of meaning" and Gibson's "affordances", Sanders notes that both thinkers believed that:

> … progress in understanding perception could be made only if philosophical bifurcations between perceiver and perceived, between "subject" and "object," could be avoided. (1993, p. 289)

Sanders suggests that their approaches were complementary and, although the "metaphysical/epistemological implications were almost never explicit" in Gibson, his "contributions are of vital importance to the deeper projects that engage Merleau-Ponty". And, particularly, with regard to affordances:

> The Gibsonian notion of affordances encapsulates in a single idea both the necessarily perspectival character, and the primitive meaningfulness of the world, and thus makes its contribution to the programme that is most closely associated with Merleau-Ponty. (Sanders, 1993, p. 298)[124]

Other authors draw similar parallels between Gibson's ecological approach and the phenomenological tradition, eg Clark (1997, p. 172) , Dreyfus (2005, p. 12), Kelly (2005, p. 102). Chemero draws parallels with the radical empiricism of William James (2009, pp. 140, 183)[125], and describes "Gibsonian ecological psychology" as a "direct descendent of the work of

---

[124] And in a later paper Sanders discusses the ontological issues that arise from Gibson's approach to perception and concludes: "Thus the ontological basis for the ecological approach cannot be any version of materialism or idealism. And what's most important about all of this is that affordances themselves seem like excellent candidates not only for primitive terms in the explanation of perception, but as primitive to the deeper task of ontology itself." (Sanders, 1997, p. 111)

[125] Here Chemero discusses "Heft's casting of Gibson as a radical empiricist and neutral monist. Neutral monism is the view that mind and the world are composed of "pure experience", which is in itself neither mental nor physical." (Chemero, 2009, p. 183)

James and Dewey." (p. 30). Oytam and Neilson (2007) also cite James as influential but major on the connection between Gibson's "ecological realism" and Berkeley's idealism.

Anti-realist associations are to be found in Gibson's intellectual successors. Chemero (2009, p. 184) cites several authors[126] who have made the "case that embodied cognitive science, radical or not, is out of line with metaphysical and scientific realism." The connection with anti-realism is particularly strong when, in its "radical" form, embodied cognitive science follows Gibson all the way in rejecting the notion of representation.

As Zahidi points out in his paper *Non-representationalist cognitive science and realism* (2013, p. 461), some of the pioneers of the non-representational approach claim that it is "incompatible with realism, viz. the idea that the world is mind-independent." Referencing one of the earliest and most influential texts introducing this approach, the *The Embodied Mind* (Varela, Thompson, & Rosch, 1991), Zahidi notes that these authors "seem to deny that a world exists that is independent of the organisms." Varela et al certainly appear to make this denial, and moreover make the connection between the rejection of realism and the rejection of representations. In outlining the representation-averse position they wish to defend, they reject what they see as mainstream cognitive science's "commitment to realism or objectivism/subjectivism about the way the world is, what we are, and how we come to know the world." (Varela, et al., 1991, p. 9)

Zahidi maintains, against those such as Varela et al, that "non-representationalist cognitive science can be pursued without resorting to an anti-realist metaphysical view" (2013, p. 462), and to this end he rightly rejects a number of arguments which fail to distinguish between directness and realism claims. Indeed, the literature is compromised by a failure to make this

---

[126] Varela et al (1991), Cousins (1992), Smith (1996), Chemero (1998b), Keijzer (2001), and Rhode and Stewart(2008)

distinction.[127] In our terminology, although considerations with respect to the constitution and operation of our, and other, species' perceptual systems might provide the basis of objections to the metaphysical directness-claim, they have no purchase against the realism-claim. The latter is a claim about how the world is, irrespective of whether, and how so, it is perceived. To address the connection between the rejection of representations and the realism-claim an entirely different line of argument is required. We suggest such an argument, in Appendix A. Here, in the main body of the dissertation, we are concerned only to establish a prima facie connection between direct realism's realism-claim and the debate between mainstream and Gibsonian cognitive science. To this end we make three further points.

Firstly, although Zahidi maintains that "non-representationalist cognitive science" is compatible with realism, he defends only a highly qualified version of realism. Zahidi proposes a version of Hacking's "entity realism" (Hacking, 1983) which builds on the notion of an entity being real if it can be used to manipulate, and suggests that we use manipulation as a "mark of the real". However Zahidi recognizes that since "what can be manipulated depends on an organism's body type", his manipulation thesis will provide a "relativized reality":

> More precisely, for an entity x to be real for a certain organism type Y, is according to the manipulation thesis defined by: x is real for Y if and only if Y can manipulate x.
>
> The world for an organism type Y then consists simply of all the entities that are real for Y. (Zahidi, 2013, p. 11)

He goes on to show that this relativism might be mitigated:

> Although it [reality]is a relativized with respect to an organism type, it is not relativized to individual organisms. So what is real does not depend on the individual organism but on the whole species. (p. 11)

However, he concedes that, it may be argued that although, by positing a reality relativized to the species rather than the individual we "get some form of objectivity,… it is a far cry from what is

---

[127] Indeed, it seems that Zahidi himself is not always clear on this point. And, as we will see below, even Chemero, who discusses these issues in depth, does not always maintain the distinction.

usually understood by objectivity and realism." (p. 12) Although Zahidi does attempt to address these concerns[128], questions remain and it does seem that such an "interpretation of objectivity", is indeed a "far cry" from direct realism's realism-claim.

Secondly, although Chemero, a leading proponent of representation-averse radical embodied cognitive science, no longer endorses anti-realist arguments (1998), he now (2009) aligns with Zahidi in defending a version of Hackings entity realism. Thus the same reservations arise, with respect to the robustness of such a relativized realism. Moreover, Chemero draws connections between the rejection of representations, direct perception, affordances and "ontological funny business" (2009, p. 135) which appears to be incompatible with the realism-claim.

Thirdly, we find the same connection between the rejection of representations and an ontology which appears to be incompatible with the realism-claim in Gibsonian Ecological Psychology. Oytam and Neilson explore the connection between "ecological realism" and "Berkeleian principles". After a detailed analysis of the ontology of "ecological realism", they conclude that:

> In ecological psychology, we have found that an anti-representational perspective of perception-action grounded upon an ontology within which the principles of [Berkeleian] idealism appears consistently to inhere. (Oytam & Neilson, 2007, p. 302)

This may be a "realism" of sorts but, here again, it is a far cry from the realism-claim of contemporary direct realist (analytic) philosophy.

3.3.4 Consolidating the alliance

Here, in Section 3.3, we have been concerned with the relationship between the science and philosophy of perception, and particularly between the "direct perception" rejected within mainstream cognitive science and the "direct realism" espoused within contemporary analytic

---

[128] At this point Zahidi questions the skeptic's assumption that "a defence of realism has to be couched in terms of a fundamental, underlying real world", and references "alternative versions of realism in which reality is viewed as a patchwork of worlds." Cartwright (1999) and Dupre (1995) "defend such a non-fundamentalist, i.e. non-reductionist, view of scientific theories." (Zahidi, 2013, p. 12)

philosophy. We noted at the outset that opinions divided as to the relevance of the one to the other. Moreover, as we saw in Chapter 2, although rarely explicit in the recent literature, an overarching irrelevance-objection could be voiced against the argument from causation: that considerations with respect to the physical and physiological mechanism of perception are, in principle, irrelevant to the thesis of "direct realism" which is being defended. We agreed that "direct realism" could be defined in such terms, but maintained that such a deflationary construal, as a *strictly-conceptual* thesis, would not do justice to the thesis which has been at issue since the Early Moderns, was largely rejected until the middle of the last century, and is now defended by contemporary direct realists.

We can now go further to maintain that any such deflationary construal would fail to address the issues which arise in the debate between Helmholtzian and Gibsonian cognitive science. Indeed it is difficult to see how an examination of "direct realism", construed as a significant and substantial thesis about this world in which we find ourselves and our place in it, could stand apart from the developments with respect to embodiment, computation and representation discussed in Section 3.2, or the realism debate outlined in Section 3.3.3.

Again, it must be emphasized, that we do not deny that "direct realism" could be construed as a *strictly-conceptual* thesis which stands apart from any such *factual* considerations. Indeed such a construal was de rigueur in the middle of the last century. The key point is that such a construal would not do justice to the thesis which is defended by contemporary direct realists, would not address the issues that stand between Helmholtzian and Gibsonian cognitive science and, a fortiori, would not address the issues at the heart of the "puzzle of experience" (Valberg, 1992) or "one's conception of one's place in the natural world." (Noe & Thompson, 2002)

Against any such demarcation, we maintain that the significant and substantial issues which stand between theories of direct and indirect perception can only be progressed by an alliance

between philosophy and cognitive science. Indeed, as we noted in Section 1.2.3 such an alliance was never questioned until the middle of the twentieth century when, with Linguistic Philosophy in the ascendant, Ryle was reminding fellow philosophers that they should stick to their last and that what they were "after is [only] an account of how certain words work". (Ryle, 1953, p. 185).

We concur with Hirst who, with Ryle in his sights, takes issue with those who seek to divorce the philosophy of perception from the empirical science. He argues that philosophy should continue in the tradition of Descartes, Locke, Berkeley, Hume and Kant to:

> … press into philosophical service the discoveries of science in order to achieve the aim of a full understanding of the nature of man and his relation to the world. That it is no longer fashionable to regard such an aim as philosophical is a departure from the tradition, and it is difficult to see it as a cause for rejoicing or self-congratulation. (Hirst, 1959, p. 135)

Similarly, again in the context of "the problem of perception" and with the same targets in view, Pasch writes:

> On the assumption[129], which I freely endorse, that it is wrong for the philosopher to arbitrarily cut himself off from any source of knowledge, the onus of the argument in a case like this is on whoever would deny the philosophic significance of scientific evidence and theories about perception. (1957, p. 55)

Five decades on, with the advances in the scope and sophistication of cognitive science, the onus of the argument is still more strongly focused on those who would make this denial. Indeed, in the face of the discussion in Sections 3.2 and 3.3, it would be very difficult to claim that the debate within cognitive science has no "philosophic significance".

Churchland's *Brain-Wise Studies in Neurophilosphy*, serves to demonstrate the weight of the opposition to any such demarcation. This monograph is divided into three sections[130] the most

---

[129] Pasch supports his "assumption" by raising and the rejecting five arguments that might be taken to support the separation between the philosophy and science of perception. Arguments 1 and 2 concern the epistemological foundations of science, argument 3 maintains that "science …. is concerned with *how* we perceive, but the epistemological problem is *what* we perceive", and arguments 4(a) and (b) concern the distinction between factual and respectively linguistic and conceptual claims.

[130] The third is a relatively brief discussion of the implications for religion.

substantial of which, headed "Metaphysics" and "Epistemology" respectively, makes the case

with detailed examples and discussion that certain findings of neuroscience are directly relevant

to the very questions that have exercised philosophers over the last two millennia. Importantly,

Churchland is not arguing that neuroscience should take over the role of philosophy, but rather,

adding weight to Hirst and Pasch in arguing against those philosophers who, "wish to isolate

philosophy from science." (2002, p. 32)[131]

In an interview, two years after its publication, she explains her motivation for writing *Brain-Wise*:

> I wanted to show that much of what academic philosophy continues to honor as "conceptual analysis" is actually hocus-pocus: it advertises itself as mere clarification of concepts, but it secretly winches its way along by theorizing about the nature of things, such as consciousness, knowledge, and free will. When challenged with the facts from neuroscience, it says, "well, conceptual analysis is just about meaning, not facts". When challenged that the analyses do not correspond to recognizable common meaning, it says "well, this is about what is conceptually necessary, and we explore that with 'thought experiments'." But what anyone thinks is conceptually necessary – and the form his thought experiments take – depend crucially on what he believes are the facts of the matter. (Churchland, 2004)

Here is the same criticism as that made by Pasch of those isolationist philosophers who claim

that "whereas the scientist is engaged in an investigation of the world, the epistemologist[132] is

engaged in an analysis of concepts." Churchland charges the philosopher, who claims to be

concerned only with "concepts" rather than "facts", with "secretly ….. theorizing about the

nature of things" (2004), whilst for Pasch the charge is having "one eye on the facts" (1957, p.

58) and "smuggling in empirical evidence" (p. 56). Importantly, neither Churchland nor Pasch

---

[131] Churchland's challenge is not just to philosophers, but also to cognitive scientists who avoid brain science. "Neurophilosophy arises out of the recognition that at long last, the brain sciences and their adjunct technology are sufficiently advanced that real progress can be made in understanding the mind-brain . More brashly, it predicts that philosophy of mind conducted with no understanding of neurons and the brain is likely to be sterile." (Churchland, 2002, p. 3)

[132] Here Pasch appears to be using the term "epistemologist" to refer to refer to philosophers in general, rather than to draw a distinction between epistemology and metaphysics.

claims that the analysis of concepts is unnecessary, but rather that such an analysis cannot be conducted in isolation from the relevant facts. In *Brain-Wise*, Churchland cites approvingly Quine's challenge to the analytic/synthetic distinction, but notes that "few philosophers took his conclusions seriously enough to change the way they did business." (2002, p. 268) Smythies puts it more strongly, and refers disparagingly to those direct realists who claim that they are "dealing with the 'logic' of perception which leads a life miraculously independent from the scientific account of how perception actually works." (1993, p. 218)

In the context of our examination of the argument from causation, the danger of attempting conceptual analysis in isolation from the relevant facts is exemplified by the widespread assumption of absolute simultaneity in discussions of the time-gap and the disappearing star. Although, from the perspective of our pre-critical intuitions and Newtonian physics, this notion might have appeared to be a conceptual necessity, it is now, in accordance with Einstein's 1905 theory of Special Relativity, rejected as factually incorrect. There is in fact no absolute simultaneity in this our actual world. (Feynman, 1963, pp. 15-6)

The necessity of both the analysis of concepts and empirical investigation is best exemplified with respect to the notion of representation, which will occupy the remaining chapters of this dissertation. Ramsey questions the contribution of the philosopher, and answers:

> … that in the cross-disciplinary enterprise of cognitive science, what philosophers bring to the table is a historical understanding of the key notions like representation, along with the analytic tools to point out the relevant distinctions, clarifications, implications, and contradictions that are necessary to evaluate the way this notion is used (and ought not to be used). To some degree our current understanding of representation in cognitive sciences is in the state of disarray, without any consensus on the different ways the notion is employed, on what distinguishes the representational theory from a non-representational one, or even on what something is supposed to be doing when it functions as a representation." (Ramsey, 2007, p. 7)

Similarly we might question the contribution of the scientist, and answer that he/she brings to the table a detailed understanding of how representations can be instantiated in rule-based symbol-systems, connectionist networks and dynamical systems, and how such representations might provide an explanation of the efficacy of perceptually guided behaviour and the richness of perceptual experience. Without such an understanding, any discussion of representation in cognition will be at best superficial and seriously compromised by a failure to appreciate, with reference to fleshed-out examples thereof, what representations actually are, what they represent and how they represent.

Perhaps it might be useful to think of "philosophers" and "cognitive scientists" as bringing particular skills and specialism's to the cross-disciplinary enterprise of perception studies. However, in many cases, even this pragmatic distinction would be difficult to sustain into the twenty first century. Where would we place, for example, the contributions of Wheeler (2005), Chemero (2009), Shapiro (2011) and Churchland (2012)?

From this point on in the dissertation, any distinction between the "science" and the "philosophy" of perception will fade into the background. Chapter 4 (and Appendix A) will, as advertised in the introduction, concern philosophically sophisticated cognitive science, or scientifically savvy philosophy, call it what you may.

3.4 Conclusions and way forward

1. Although direct realism is now the orthodoxy within much of contemporary analytic philosophy, the claim that perception is direct is still rejected within mainstream cognitive science; here the tables are turned and direct perception remains the heterodoxy.

2. Gibson's avowedly revolutionary approach to perception (1979) presented a challenge to the Helmholtzian orthodoxy, albeit at the expense of rejecting methods and assumptions that had previously gone unquestioned. Although Gibson's theory of direct perception remains highly controversial, it has been developed by an influential minority, notably proponents of radical embodied cognitive science, who have drawn support from recent work on embodiment, computation, and representation. We do not take sides in the debate. The key point is that, when theories of direct perception are maintained, it is at the expense of rejecting much of orthodox cognitive science and espousing a radical alternative which is at odds with the standard causal picture.

3. The debate within analytic philosophy cannot be divorced from that within cognitive science. That perception involves inference, is a key tenet of Helmholtzian cognitive science, whereas Gibson's opposition to the involvement of inference aligns with the epistemological directness-claim. With respect to the metaphysical directness-claim, Helmholtzian cognitive science affirms the standard causal picture and thus endorses the argument from causation, though the extent to which it advances the traditional form of the argument is open to question. On the other hand by rejecting the standard causal picture, and thus disarming the argument from causation, Gibsonian cognitive science lends support to the metaphysical directness-claim.

4. Although the realism-claim is unquestioned within mainstream cognitive science and direct realist philosophy, it does become an issue in the theories of direct perception proposed by Gibson and his intellectual successors. Here representation is the pivotal issue. There is, at least, a prima facie connection between representation-averse theories of perception and

ontologies which compromise the realism-claim. Indeed such theories have been associated with the Phenomenological Tradition, Berkeleian idealism and Jamesian radical empiricism.

5. At the end of Chapter 2 we concluded that although our argument from causation was robust, and could not be dismissed lightly, two objections might be raised: the type-2 overarching irrelevance objection and the type-3 objection concerning the notion of an immediate object of awareness. Against the former, we argued that although "direct realism" could be defended as a *strictly-conceptual* thesis which stands independently of any considerations with respect to the physical and physiological mechanism of perception, such a deflationary construal would not do justice to the thesis which is actually defended by contemporary direct realists. Drawing on the debate between Helmholtzian and Gibsonian cognitive science, and recent developments with respect to embodiment, computation and representation, we can go further. Such a deflationary construal of "direct realism", would have no point of contact with these developments and would, in comparison to the Gibsonian notion of direct perception, constitute an insubstantial and insignificant thesis. A fortiori, the deflationary construal as a *strictly-conceptual* thesis does not address the issues at the heart of the "puzzle of experience" (Valberg, 1992) or "one's conception of one's place in the natural world." (Noë & Thompson, 2002)

6. Having disposed of the overarching irrelevance objection, only type-3 objections remain to be addressed. However as we have seen in Chapter 2 such objections can be raised not only by those thoroughgoing reductive representationalists who reject the assertion of the Weak Phenomenal Principle at Premise 2, but by others who might question the arguments call on the notion of an immediate object of awareness at premises 6 and 7. To break this impasse we will need to go beyond the standard causal pictured to formulate an argument against

representationalist direct realism, which does not call on the notion of an immediate object of awareness. To this end, in Chapter 4, we will draw on the representations posited within cognitive science to examine more closely the viability of representation-friendly direct realism.

7. Although we accept that Gibsonian cognitive science disarms the traditional version of the argument from causation which takes the standard causal picture as a premise, it does not threaten our conjunctive version which claims only that the metaphysical directness-claim and the standard causal picture cannot both be true. However, it could be maintained that by virtue of positing a non-standard, but nevertheless naturalistic, causal picture, it does present an objection to our central thesis. We would reject this suggestion on two counts. Firstly there are, as yet, no fleshed-out representation-averse "explanations", and arguably none in prospect; according to mainstream cognitive science, perceptually guided behaviour can only be explained in terms of representations.   Secondly, although Gibsonian cognitive science might be compatible with the directness-claim, it is not robustly realist; indeed some of its proponents are avowedly antirealist.

 Although, outside the scope of this dissertation, we would go further to argue that representation-averse cognitive science is, in principle, incompatible with the realism-claim. We suggest, as the basis for further work, such an argument in Appendix A. The argumentation, which is inevitably less direct and the conclusions more tentative than those in Chapter 4, proceeds in two stages: firstly that perceptually guided behaviour can only be explained without recourse to representations by positing a new ontology and, secondly, that such an ontology is incompatible with the realism-claim.

Chapter 4: Representation-friendly cognitive science and the metaphysical directness-claim

As we saw in Chapter 3, according to mainstream cognitive science perception involves inner representations of the external world and, for this and other reasons, perception is taken to be indirect. However, with respect to direct realism, the notion of representation cuts both ways. Negatively, representations can be seen as the perceptual intermediaries posited by indirect realism, but on the other hand the notion of representational content is essential to the representationalist/ intentionalist version of direct realism. Indeed, as discussed in Section 2.2.3, proponents of Strong Representationalism account for the immediacy of perception in terms of an identity between "phenomenal character" and "representational content" (Tye, 2009, p. 256). Moreover, by virtue of rejecting the Weak Phenomenal Principle, such a reductive representationalist version of direct realism presents the one remaining objection to the argument from causation formalised in Section 1.3.

As advertised in Section 3.3.2, to break the impasse we need to go beyond the mere involvement of representations, to understand what representations actually are, what they represent and how they represent. We need to know whether any of the various representations posited by cognitive science could do the work required of a representation in a direct realist theory of perception.

In Section 4.1 we examine the role of representations in such a theory. Here we introduce the key distinction between the work required of a representation as the supplier of content in representationalist direct realism, role A, and in an explanation of perceptually guided behaviour, role B.

In Section 4.2, before addressing the particular subtleties of representations in cognition, we ground our discussion on the simulation models of the physical world widely used in science and engineering. Here representations are instantiated in the form of data structures in rule-based

symbol-systems which are second-order isomorphic to their targets, and their efficacy in predicting the behaviour of the systems they represent is explained in terms of surrogate reasoning. We conclude that such representations, though well suited to role B, could not supply the representational content required in role A. The question is now whether any of the representations posited by cognitive science are more suited to role A.

Drawing particularly on the work of Marr (1982) and Churchland (2002), Sections 4.3 and 4.4 examine representations in orthodox and connectionist cognitive science. Here again the representations are found to be second-order isomorphic to their targets and, as such, do not have the resources to meet the requirements of role A. In Section 4.5 we consolidate our argument and extend the scope of our conclusions. Section 4.6 goes beyond the defence of our central thesis to question, in the context of Locke, Berkeley and the Churchland's "slippery slope" (2002) (2012), the coherence of the theories of perception posited within representation-friendly cognitive science; although beyond our remit, it would be a significant omission not to voice these concerns. Section 4.7 summarizes our conclusions.

4.1 The role of representations

Any representation-friendly theory of perception needs to address two issues: the relationship between the world and the representation thereof, and the relationship between the representation and the phenomenal character of perceptual experience. Moreover, a direct realist theory of perception needs to address both issues in a way which is compatible with direct realism's directness and realism claims.

The representationalist theory of consciousness addresses the second issue in a manner which, its proponents maintain, provides both a defence of direct realism against the arguments from illusion and hallucination and an accommodation of qualia within a materialist theory of mind. (Lycan, 2015) Many variants of the representationalist theory of consciousness have been

discussed in the recent literature[133], but they agree on the essential point that the phenomenal

character of our experiences is determined by their representational content. Tye, a proponent of

the representationalist theory defines "Strong" and "Weak" Representationalism in terms of

identity and supervenience respectively:

> *Strong Representationalism*: phenomenal character is one and the same as representational content that meets certain further conditions.
>
> *Weak Representationalism*: phenomenal character supervenes on representational content that meets certain further conditions…. (2009, p. 256)[134]

For the purpose of this discussion we will grant that Strong Representationalism is true and that

the "certain further conditions" are met. By being, in this respect, charitable to the

representationalists we can put aside all issues relating to the genesis and nature of qualia.

However the first issue, the relationship between the world and the representation thereof,

remains to be addressed and is the focus of our critique of representationalist direct realism.

It is important to emphasize that Strong Representationalism only addresses the relationship

between the representation and the perceptual experience and does not, on its own, constitute a

direct realist theory of perception. The representationalist version of direct realism also requires

that the relationship between the world and the representation thereof is such that the

representation can supply the necessary representational content. Representationalist direct

realism requires a representationalist theory of consciousness, plus a representation able to

---

[133] They include: natural and conventional (Dretske, 1997) intermodal and intramodal, and restricted and unrestricted (Byrne, 2001); realist and anti-realist, content-based and vehicle based, and ontologically liberal and conservative (Thompson, 2008); strong and weak, and reductive and non-reductive (Tye, 2009); and narrow and wide (Lycan, 2015).

[134] Kind puts the point well: "According to representationalism, the qualitative character of our phenomenal mental states supervenes on the intentional content of such states. Strong representationalism makes a further claim: the qualitative character of our phenomenal mental states *consists in* the intentional content of such states." (Kind, 2007, p. 405)

supply the necessary representational content. Provisionally, we can outline the job description[135] of such a representation as:

Job description role A: *To supply the representational content needed to meet the requirements of the representionalist version of direct realism.*

However it is essential to distinguish role A from the more usual role of representations in cognitive science, that is in explaining perceptually guided behaviour. The job description for representations in this second role can be outlined as:

Job description role B: *To supply the representational content needed to explain perceptually guided behaviour.*

The distinction between the two roles needs to be emphasized. The requirement for a representation in role B is to provide the basis for an explanation of perceptually guided behaviour[136], that is behaviour that could be studied by a third party, say an experimental psychologist or roboticist. The reference to robotics is significant, since the success criteria for a representation in role B is often taken to be the demonstration of perceptually guided discriminations in non-human mechanisms. The crucial point is that neither the mechanisms nor the success criteria need make reference to perceptual experience. Work on representations in role B could equally well take place in the context of zombie[137] perception; nothing would be lost.

---

[135] We take the notion a "job description" in the discussion of representations from Ramsey (2007, p. 24), though we develop the idea in a different direction.

[136] The emphasis on the role of representations in the control of behaviour is evident in Haugeland's introduction to his much cited *Representational Genera*: "A sophisticated system (organism) designed (evolved) to maximise some end (such as survival) must in general adjust its **behaviour** to specific features, structures, or configurations of its environment …. But if the relevant features are not always present (manifest), then they can, at least in some cases, be represented; that is, something else can stand for them with the power to guide **behaviour** in their stead [emphasis added]." (1990, p. 62)

[137] Here the term "zombie" is used in the technical sense. (Kirk, 2015)

In contrast, the requirement for representations in role A is to supply the representational content

necessary to account for the phenomenal character of perceptual experience. Fish describes well

just how demanding this requirement will be:

> [Since] the representationalist insists that what it is like to have an experience is
> explained by its representational content, an adequate theory of content must ascribe
> contents to experience that are as rich and detailed as that experience's phenomenology.
> (2010, p. 78)

Moreover, in role A, a representation must be able to supply representational content such that

the perceptual experience is not only sufficiently rich and detailed, but is also in accordance with

the metaphysical directness-claim. This claim can be expressed in a number of forms, but for our

present purposes, that is a critique of representationalist direct realism, it is desirable to use a

form of words which avoids reference to an "immediate object of awareness" and has the

imprimatur of the representationalist camp. To this end, rather than the wording introduced in

Section 1.2.2[138], we draw on the reference to "direct realism" in the *Oxford Companion to*

Philosophy which directs the reader to Dretske's entry on "naïve realism"[139] which is described

as:

> A theory of perception that holds that our ordinary perception of physical objects is
> direct, unmediated by awareness of subjective entities, and that, in normal perceptual
> conditions these objects have the properties they appear to have. (1995, p. 602)

The metaphysical directness-claim has both a negative and positive component. Negatively, the

claim is that there are no perceptual intermediaries. Positively, the claim is that in veridical

perception "[physical] objects have the properties that they appear to have". Thus, the

---

[138] In Chapter 1, taking the lead from Smith, we introduced the metaphysical directness-claim to be as an identity claim: "the immediate object of awareness in standard perceptual situations is a normal physical object - in other words, that it is *identical* to some such object." (2002, p. 8) However, in this form, the directness-claim is open to the representationalist's objection to the notion of an "immediate object of awareness". In Chapter 1 we addressed this objection by drawing on the Weak Phenomenal Principle. Nevertheless, in the context of a critique of representationalist direct realism, it is preferable to express the directness-claim without introducing the contentious notion of an immediate object of awareness.

[139] Dretske takes the terms to be synonymous.

requirement on a representation in role A is to supply the representational content in accordance with both the positive and negative components of the directness-claim.

With regard to the negative component, we will accept that the role of the representation in the perceptual process is such that it (the representation) does not constitute a perceptual intermediary. This concession is implicit in our acceptance of Strong Representationalism and without it (the concession) representationalist direct realism does not get off the ground. We will focus our critique on the positive component.

The positive component of the directness-claim, that "[physical] objects have the properties that they appear to have" implies that (i) objects have properties, (ii) objects appear to have properties and (iii) in veridical perception objects appear to have the properties that they have. Thus, if, as the representationalist claims, objects appear to have the properties they are represented as having, then by virtue of (iii), in veridical perception objects must be represented as having the properties that they have. On this basis we can refine our original job description for a representation in role A as:

Job description role A: *To supply representational content which represents objects as having the properties that they (the objects) actually have.*

The essential connection between experience, representational content and an object's properties is explicit in Dretske's exposition of the "representational view" in his paper *Experience and Representation*:

> According to a representational view of experience the phenomenal character of our experience is determined … by the way the experience represents things to be, the properties it represents objects …to have. [caveats with respect to non veridical perception omitted] (2003, p. 69)[140]

---

[140] See also the introduction to Dretske's *Naturalizing the Mind*. "If, in accordance with the Representational Thesis, we think of all mental facts as representational facts, the quality of experience, how things seem to us at the sensory level, is constituted by the properties things are represented as having. My experience of an object is the totality of

Implicit in both Dretske's definition of direct/naïve realism and our refined job description, is the assumption that objects have certain determinate properties independently of our perception of them (the objects). This is in effect the realism-claim (see Section 1.2.1) which is common to both direct and indirect realism. Role A relates specifically to direct realism since it places a particular requirement on how the properties of objects are represented. It requires that the representational content represents objects as having the properties that they (the objects) actually have rather than, as would be sufficient for indirect realism, as having properties which in some sense correspond to the properties that the objects actually have.

The role A requirement, that objects are represented as having the properties that they (the objects) actually have, is also essential to support Armstrong's claim that, in accordance with "common sense", "The world is much as it looks to be, feels to be, smells to be, tastes to be or sounds to be." (1961, p. 139) Without this positive claim, "direct realism" would be reduced to a pale shadow of the bold thesis defending our pre-critical intuitions about the nature of perception, largely rejected until the middle of the last century, defended within contemporary analytic philosophy, but still rejected by mainstream cognitive science.

The key issue is now whether or not any of the various representations posited by cognitive science, or developments thereof, can meet the requirements of role A. If not, we will have good reason to question the representationalist version of direct realism's challenge to our argument from causation. As Dretske, a staunch advocate of the representationalist theory of consciousness, notes, the theory "never really gets out of the gate" unless we have a "plausible theory of original or natural representation that can put some flesh on these bones" (2003, p. 74)

---

ways that object appears to me, and the way an object appears to me is the way my senses represent it." (Dretske, 1997, p. 1)

Given the enormous volume of literature on representation and representations, it is useful to narrow the focus of our investigations by identifying two substantial and much discussed issues with which we do not need to engage. Firstly, as noted above, we grant that Strong Representationalism is true; that is "phenomenal character is one and the same as representational content that meets further conditions" (Tye, 2009, p. 256) and, moreover, that all the cases under consideration satisfy the "further conditions". In this way we avoid altogether the question of what it is "about certain representations that makes the systems in which they occur *conscious* of what is being represented." (Dretske, 1997, p. 6) Our concern is not whether, say by virtue of teleology or natural selection, the representations are such that a system could be conscious of what is being represented but, rather, with the constraints on the representational content that the representations posited within a naturalistic theory of perception can supply.

Secondly, we do not need to address the question of what counts as a representation proper, on which, say, Ramsey (2007) and Cummins (1996) disagree with Dretske (1997) and Grush (2008). The former take a more restrictive view and would reject, for example, Dretske's notion of representation by virtue of reliable indication. Since we are concerned only with the ability of so-called representations to do a particular job, that is to meet the requirements of role A, we do not need to engage in this debate.

Moreover, although we will draw on the work that has been done on representation in the abstract, we concentrate our attention on three particular concrete examples of representations: the first in engineering, and the second and third in the theories of visual perception posited by Marr (1982) and Churchland (2002). In all three cases, we focus on the key question as to whether such representations can supply the representational content required of a representation in a direct realist theory of perception.

It might be objected that the required representations are to be found at the personal level, and that there are fairly detailed philosophical theories of how personal-level states get their contents which we do not consider. However, such theories, eg PANIC (Tye, 1997) and Representational Naturalism (Dretske, 1988) (1997), are framed at too high a level of abstraction to address the question at it issue: that is the nature of the relationship between the world and representation thereof and the consequent constraints on the type of representational content that can be provided. To address this question, we need to go beyond a consideration of representations in the abstract to examine what representations posited by cognitive science actually are, what they represent and how they represent.

This distinction is brought out in Rupert's paper *Causal Theories of Mental Content* in which he examines how such causal theories (CTs) "ground certain aspects of a concept's meaning in the causal relations a concept bears to what it represents". For the purpose of evaluating such theories he identifies two criteria: a CT must "explain misrepresentation", and a CT must "assign *correct* extensions to mental representations." However, before going on to survey various CTs[141], he notes that:

> Application of only these two evaluative criteria substantially limits the discussion. Nevertheless, it fairly captures the dimensions on which CTs have typically been evaluated, even if a genuinely satisfactory CT must meet further demands as well. (Rupert, 2008, p. 357)

And tellingly, in a footnote regarding these further demands:

> Lastly, the scientifically minded philosopher might think that a satisfactory CT must meet the demands of successful cognitive science; in other words a philosophical theory of mental content had better deliver something that can play the role asked for it by successful sciences of the mind (Cummins, *Meaning*).[142]

---

[141] Including those of Dretske, Fodor, Millikan, Papineau, Stampe and Ryder.

[142] Rupert's reference to Cummins' *Meaning and Mental Representation* (1991)is significant. Here (as advertised on the back cover) Cummins "takes on philosophers … who pursue the question of mental representation in the abstract, apart from the constraints of any particular [eg orthodox computational or connectionist] theory or framework."

We agree that any genuinely satisfactory theory of mental content must meet these demands. Moreover, a theory which is compatible with the metaphysical directness-claim must meet these demands in a manner such that representations can meet the particular demands of Role A: to supply representational content which represents objects as having the properties that they (the objects) actually have. The question of whether or not a CT can meet the requirements of Role A cannot be assessed on the basis of abstract, instantiation independent, requirements alone; this is not to suggest that the higher level considerations are unimportant.

The need to underpin their abstract theories of representation with an explanation of the mechanisms at the level of instantiations in rule-based symbol-systems or connectionist networks does not go unrecognized by their proponents. For example, Dretske presents his theory of "Representational Naturalism" at the abstract level of "conventional" and "natural" "indicator functions" (1997, p. 19). However, he accepts that without "some flesh on these bones" to explain "how biological systems actually perform this wondrous feat" such a theory "never really gets out of the gate." (2003, p. 74) Moreover he recognizes that "we need to know what kind of representation the sensory representation is" and suggests that perhaps it is "like what Marr (1982) and his associates call a 2½-D sketch?" (2004, p. 282) Similarly, Tye's PANIC theory is framed at the abstract level at which "phenomenal character" is identified with "Poised, Abstract, Nonconceptual Intentional Content". (Tye, 1997, p. 137) However, he recognizes that the high level theory needs to be underpinned by an explanation of the mechanisms involved, and here again Marr's 2 ½ D sketch is suggested as a candidate. (p. 210) (We examine Tye's references to Marr in Section 4.3.)

The key point, in defence of our focus on Marr, Churchland and Edelman in the following sections, is that our question with respect to role A must be addressed at the level of concrete instantiations, that is at the level of representations instantiated in rule-based symbol-systems and connectionist networks, rather than at the abstract level of, say, PANIC and Representational

Naturalism. The theories discussed by Marr and Churchland, are not in competition with those presented by Tye and Dretske, rather, they are framed at different levels. And, for the purposes of assessing candidate representations for role A, the critical issues only come to the fore at the lower level.

4.2 Representation in an engineering simulation

The question of whether, and how so, representations are involved in perception is problematic on two counts: firstly with respect to the nature of representation itself and, secondly, the further difficulties which arise in the context of human cognition. In this section we ground our discussion by focusing on the former in the context of the tried and tested representations of the layout of surfaces in the physical world which are widely used in science and engineering. Here we have a concrete example of a representation which we can dismantle and examine in detail. As Dretske puts it in titling a paper on the subject of representation, "If You Can't Make One, You Don't Know How It Works". (2000, p. 208) With this maxim in mind we delve into the nuts and bolts of representations.

Computer programs, known as "simulation models", within which data structures instantiated in von Neumann computers (PCs or mainframes) represent certain features of the world, are used extensively in science and engineering to predict the behaviour of complex systems. Here we consider a highly simplified example drawn from aerodynamics.

Figure 3a

Figure 3b

The airflow over even simple three dimensional structures can be difficult to analyze. However the aerodynamic drag on a complex body shape can be predicted using simulation models in
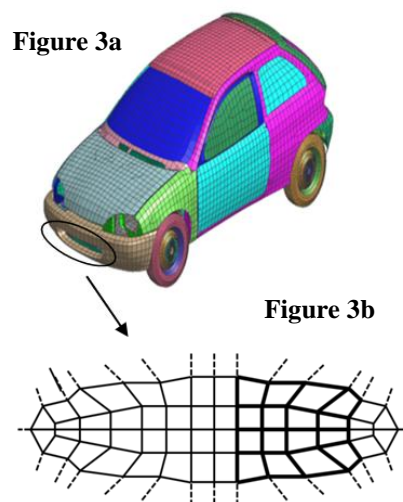
which the spatial features of the vehicle are represented by data structures[143], and the mathematical relationships between pressure and air flow are built into the algorithms which operate on the data contained therein. In such models, the curved surface of the vehicle is broken down into a large number of plane surfaces, quadrilaterals in figure 3a, which are amenable to a relatively simple mathematical treatment[144], and the total drag on the vehicle is calculated by summing the contribution from each of the quadrilateral surface elements. (It is useful to imagine that the vehicle body is actually constructed from a large number of quadrilateral shaped plates.) In this simplified example we will consider a simulation model set up to predict the drag on the forward facing surfaces of the vehicle, some of which are illustrated in figure 3b. For our purposes we are concerned only with the manner in which certain spatial features of the vehicle are represented; the physics and the algorithms are irrelevant.

It is important to note that this is just one of many possible representations of the vehicle. Different features might be represented for different purposes. For example, for the purpose of structural analysis it would be necessary to represent not only the orientation of each surface element but also its mechanical properties in terms of a number of parameters. We focus on a representation of the layout of the vehicle's surfaces, since this will align closely with representations that have been posited in theories of visual perception, notably Marr's 2½-D sketch (1982, p. 268) which is examined in the following section.

---

[143] "A data structure is a structured set of variables associated with one another in different ways, cooperatively defining components in the system and capable of being operated upon in the program." (Kamthane, 2007)

[144] The force on a plane surface can be calculated simply from its area and the velocity (both are vector quantities) of the impacting air flow. However considerable computing power and design expertise is required, since each element of the surface not only experiences a force, but also affects the velocity of the air impacting on its neighbours. Clearly there is a trade-off between the number of elements and computing power: the more elements, the finer the representation and the better the accuracy of the prediction, but the more computations required.

Consider first the representation of a single quadrilateral surface element. To specify the shape and orientation of a quadrilateral in three dimensional space, we will assume[145] that four sets of three dimensional Cartesian coordinates (x, y, z) are needed, one set for each vertex. The simplest data structure by which to represent such a quadrilateral is a 6-column by 2-row, two dimensional array, in which the coordinates of each vertex of the quadrilateral are positioned correspondingly in the array. See figure 3c. One such array will be required to represent each quadrilateral element of the vehicles surface.
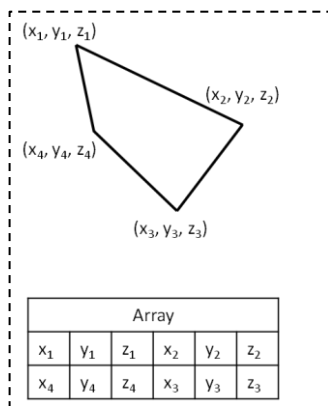

Figure 3 c

---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
---	---	150	310	222	210	300	220	270	305	200	350	290	216	370	270	223	---	---
---	---	150	250	210	200	250	210	250	250	210	300	240	220	360	200	230	---	---
---	---	150	200	210	200	200	210	250	200	210	300	200	220	360	200	230	---	---
---	---	150	150	210	200	150	210	250	150	210	300	160	220	360	170	230	---	---
---	---	150	90	222	210	100	220	270	95	200	350	110	216	370	130	223	---	---
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Figure 3d

Since we have restricted the tessellation of the vehicle's surface, such that each quadrilateral must have 8 neighbours[146], the adjacency relationships between the quadrilaterals can be represented straightforwardly in a regular data structure which brings together the arrays as shown in figure 3d. Here we can see how the 6 ×2, 12-element, array representing each of the quadrilaterals is positioned within the data structure, which is itself a larger two dimensional array. The numerical values in the data structure, figure 3d, represent the bolded surfaces in figure 3b, and for the purpose of the following discussion we need to assume that the other surfaces are similarly represented in the data structure. Thus, we will suppose that figure 3d is a data structure representing the complete forward facing surface of the vehicle.

---

[145] This is not strictly true. Since all four points must lie on a plane, only 11 coordinates needed. Thus the explicit representation of all 12 coordinates is a trade-off between information accessibility and redundancy.

[146] Diagonals included.

The first point to note is that the data structure is the representation of the vehicle; so here we have a real concrete example of a representation.[147] Figure 3d is not just an illustration of a representation, it is the real thing. However the same data structure, that is representation, can be instantiated in a variety of forms. Here, in printed form as figure 3d, the representation is instantiated in the form of ink on paper. But the same representation is, at the time of writing (on a laptop) also instantiated in the distribution of electric charges on silicon chips (RAM) and the orientations of magnetic domains on disk (hard drive). Although these various instantiations differ in terms of the physical medium and conventions used, they are nevertheless all instantiations of the same data structure, that is instantiations of the same representation of the vehicle. We will return to this point below. "Representation" and "representational content" may be slippery concepts, but we have here, in the form of a data structure instantiated on paper, on silicon chips and on magnetic media completely perspicuous examples of representations of the layout of surfaces in the physical world.

Note also that the representation supplies information, representational content, to a user. The paper and ink instantiation, figure 3d, might be used by an engineer to hand calculate the drag on a few quadrilateral surface elements in order to validate the logic of the simulation software. When the simulation is running on a PC or mainframe the instantiation in RAM is used by the algorithm which calculates the drag. The notion of a user of representational content will follow through into the representations posited in cognitive science.

The question now is in what sense does the data structure represent the surface layout of the vehicle. Ramsey (2007), among others, draws on Pierce's theory of semiotics to classify the

---

[147] Strictly speaking, the representation is the data structure plus the operations needed to extract the information required by the user. When the simulation model is running and the representation is instantiated in RAM, the user would be the algorithms which calculate the drag on each of the quadrilateral surfaces. As Palmer puts it: "the only information contained in a representation is that for which operations are defined." (1978, p. 264) Similarly, according to Marr: "A *representation* is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this." (1982, p. 20)

various notions of representation in terms of the relationship between the representation and its object (that which it represents)[148]. Pierce distinguished between three types of representation in terms of the different ways that the representations[149] are related to their objects: icons, indices and symbols. Icons are connected to their object "by virtue of some sort of structural similarity or isomorphism between the representation and its object." Indices represent "by virtue of some sort of causal or law-like relation" between the object and its representation. Lastly, symbols are related to their objects entirely by habit or "convention", that is the connection is purely arbitrary. (Ramsey, 2007, p. 21)

In terms of Ramsey's taxonomy, the representation in our example is of the icon type. There is a "structural similarity", not merely a "law-like relation", between the data structure and the surface of the vehicle. Consider how each surface element of the vehicle is represented. For example, quadrilateral surface elements Q1, Q2 and Q3 are represented by arrays A1, A2, and A3 as shown in figure 4. (The curved arrows serve to show that the surfaces differ not only in size and shape but also in orientation.)

Array 1					
270	305	200	350	290	216
250	250	210	300	240	220

Array 2					
200	250	210	250	250	210
200	200	210	250	200	210

Array 3					
250	200	210	300	200	220
250	150	210	300	160	220
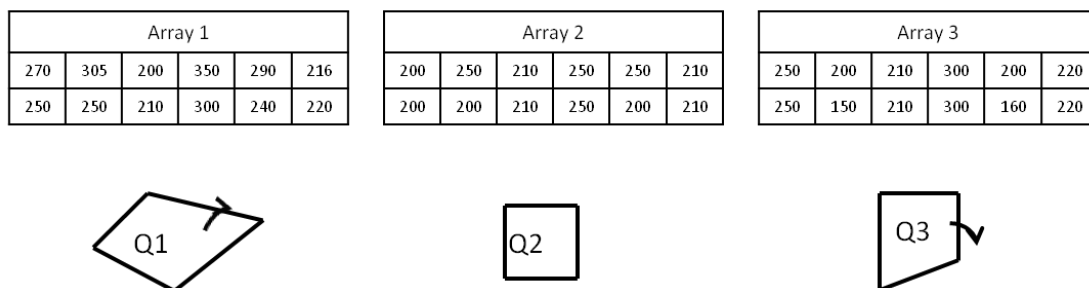
Figure 4

---

[148] The terms "object" and "target" are both used to refer to that which a representation represents. The term "target" is more neutral and to be preferred, but "object" will be retained where the author in question uses the term.

[149]Pierce talked in terms of "signs" rather than "representations".

Although there is no resemblance in a pictorial sense, the arrays are isomorphic to the

quadrilaterals in the mathematical sense.[150] That is, there is a one to one correspondence between

the quadrilaterals and the arrays, such that result of an operation on the quadrilaterals

corresponds to the results of an analogous operation on the corresponding arrays.[151] Indeed there

are two isomorphisms between the quadrilaterals and the arrays: with respect to area-type

operations and to orientation-type operations.

Consider the isomorphism with respect to area. There is a one to one correspondence between

the quadrilaterals and the arrays, such that an operation on a quadrilateral surface, that is

measuring its area (perhaps by pasting squared paper on the surfaces and then counting the

squares), corresponds to the results of an analogous operation on the corresponding array, that is

applying a particular computation[152]. In this case the particular computation would give 4000,

2500 and 2130 for A1, A2 and A3 respectively. The results of the measurement operation and its

computational analogue correspond in the sense that certain relationships between the results of

the measurements on the quadrilateral surfaces are preserved in the results of the computations

on the arrays. Clearly the ordering by area of the quadrilaterals, Q1, Q2, Q3 (largest to smallest),

is preserved in the results of the computations on the corresponding arrays, 4000, 2500, 2130.

Less obvious, but essential to the simulation, is the fact that not only ordering but also proportion

---

[150] The term "isomorphism" literally means equality or sameness (iso) of form (morphism). It is widely, and somewhat differently, used in a number of disciplines including mathematics and cognitive science, but also biology, geology and sociology. In general terms the mathematical notion of isomorphism can be defined as: "A one-to-one correspondence between the elements of two sets such that the result of an operation on elements of one set corresponds to the result of the analogous operation on their images in the other set." (Dictionary, 2016)

[151] Although the distinctions will not be important in our discussion and there are useful parallels with the mathematical usage, here in cognitive science the term "isomorphism" is not used in the strict mathematical sense. In mathematics, isomorphism is a particular type of "homomorphism" which requires that the mapping between the two sets is strictly "one-to-one" (an "injection") and "onto" (a "surjection"). (Weisstein, 2013) Neither restriction applies to the usage in cognitive science. Thus in mathematics, but not in cognitive science, isomorphism is necessarily symmetric, transitive and reflexive. Moreover, in mathematics, isomorphism is by definition a second-order relationship; the distinction between first and second-order does not arise.

[152] The particular function is: $\frac{1}{2}|(V_2 - V_o) \times (V_3 - V_1)|$, where V_0 is the vector (x_0, y_0, z_0) etc, $||$ is magnitude and $\times$ the vector cross product.

is preserved. For example, both the results of area measurements on Q1 and Q2, and the results of computations on the corresponding arrays (4000 and 2500) are in the proportion 16:10. Similarly, proportion is preserved in respect of orientation measurements on the quadrilateral surface elements, and orientation-type computations on the corresponding arrays.

The key point is that, with respect to area and orientation, the similarity is not between Q1 and A1 etc, but rather the similarity is between the relationship between Q1, Q2 and Q3 etc and the relationship between A1, A2 and A3 etc. In terms of the distinction introduced by Shepard and Chipman, this is an example of "second-order isomorphism", that is:

> … the second-order relationship between (a) the relations among alternative external objects, and (b) the relations among their corresponding internal representations. (1970, p. 2)

So, in Shepard and Chipman's terminology, the arrays represent the area and orientation of the quadrilateral surface elements by virtue of "second-order isomorphism". Swoyer puts it succinctly, "the *pattern* of relations among the constituents of the represented phenomenon", here of the surface elements, "is mirrored by the pattern of relations among the constituents of the representation itself", here of the arrays. (1991, p. 452)

On the other hand, first-order isomorphism involves a more direct relationship in which the representation and its target shares one or more properties. According to Shepard and Chipman there is a:

> … first order relation between (a) and individual object, and (b) its corresponding internal representation (1970, p. 2)

In a later paper, Shepard (1975) distinguishes between "concrete" and "abstract" first-order isomorphism. As Palmer (1978) points out, "first-order isomorphism" was originally taken to cover cases in which "properties of real-world objects are retained in the internal representation of those objects" such that, for example, "representations of green things must themselves be

green, and those of square things must themselves be square". However in Shepard's later paper the definition is widened to include both "concrete first-order isomorphism", as in the example of green things where the representation and its target share physically equivalent properties, and "abstract first-order isomorphism" in which there is a similarity, but no physical resemblance, between the representation and its target. As an example of the latter isomorphism, Shepard cites a representation of a square which contains four parts, each of which corresponds to a corner of the square. (Palmer, 1978, p. 290)[153]

Considering the ink on paper instantiation of the data structure, that is figure 3d, it might be thought that there is a "concrete first-order isomorphism". There appears to be a physical resemblance. For example, if surface element A is to the right of surface element B in 1b, then the representation of A is to the right of the representation of B in figure 3d. However, the layout of the data structure in physical space in figure 3d is merely an artifact of the particular ink on paper instantiation[154]; the two dimensionality of the data structure is at the conceptual level. Physical layout has no significance with respect data structures instantiated in von Neumann

---

[153] However Palmer argues that this example is ambiguous. (Palmer, 1978, p. 291) We agree, and maintain that the dividing line between abstract first-order and second-order isomorphic representation is not clear-cut. Consider for example (i) an ink on paper map showing an outline of the UK, (ii) such a map showing isobars, and (iii) digitised versions stored on hard drive. As Obrien and Opie (2004, p. 8) put it, "Second-order resemblance comes in weaker and stronger forms." Some, perhaps Edelman (1998), might regard the representations in the engineering example and Marr's to 2½ -D sketch to be abstract first-order isomorphic to their targets. We characterise them as "stronger" second-order isomorphic representations and those in Churchland (2002) and Edelman (1998) as "weaker". This issue is discussed further in Section 4.5.

[154] Indeed there is good reason to suppose that the other instantiations of the same data structure, in RAM and hard drive, do not maintain such a layout in physical space. The actual allocation of bits to capacitive cells on the surface of the RAM or magnetic domains on the hard drive is likely to be extremely complex and dependent on the details of the hardware and low-level software of these storage devices. Moreover it is possible, that for the purposes of redundancy and error protection, the representations will be distributed rather than strictly local, in the sense discussed in Section 3.2.2.

computers though this may not necessarily be the case in connectionist networks nor, a fortiori, in brains.[155]

As we move from concrete first-order, to abstract first-order and then to second-order isomorphism, the connection between the representation and its target becomes increasingly less direct. In concrete first-order isomorphism there is a similarity of physical properties, and in abstract first-order isomorphism there is a similarity of abstract properties. However, in the second-order case there is no similarity of properties (concrete or abstract) between a particular representation and its target; the similarity is between the relationship between various targets and the relationship between the corresponding representations. Indeed, in the second-order case, the connection between representation and target might appear somewhat weak.

The question now arises, as to how seemingly tenuous second-order isomorphic representations can be so effective in the engineering simulation. The answer is that second-order isomorphism enables what Swoyer calls "surrogate reasoning". Second-order isomorphs can be used in the prediction of the behaviour of complex systems since they enable "surrogate reasoning", that is "reasoning about a structural representation in order to draw inferences about what it represents":

> … the *pattern* of relations among the constituents of the represented phenomenon is mirrored by the pattern of relation among the constituents of the representation itself. And because the arrangements of things in the representation are like shadows cast by the things they portray, we can encode information about the original situation as information about the representation. Much of this information is preserved in the inferences about the constituents of the representation, so it can be transformed back into information about the original situation. And this justifies surrogate reasoning, since if we begin with true premises about the objects of representation, our detour through the representation itself will eventually wind its way back to a true conclusion about the original object. (Swoyer, 1991, p. 452)

This process of surrogate reasoning is precisely what is going on in the engineering example.

Moreover surrogate reasoning on the basis of second-order isomorphic representations of the

---

[155] In the case of the brain there is "retinoptic mapping" by which the spatial layout of cells in the retina is to some extent maintained in the layout of corresponding cells in the lower layers of the visual cortex. (Bruce, et al., 2003, p. 50)

environment may be sufficient to explain object recognition, and perhaps to explain all of perceptually guided behaviour in humans and to demonstrate such faculties in robots. Indeed, as discussed in Sections 4.4 and 4.5, we will find reason to suppose that second-order isomorphic representations are particularly well suited to role B.

However role A is very different. It does not require that the representation serves as the basis for surrogate reasoning; rather it requires that the representation can supply representational content which represents objects as having the properties that they (the objects) actually have. Here there appears to be a mismatch between the role A requirement, that is a similarity between a target and its corresponding representation, and what second-order isomorphism can provide, that is a similarity in the relationship between various targets and the relationship between their corresponding representations. As Edelman suggests, first and second-order isomorphisms are, respectively, "representation *by* similarity" and "representation *of* similarity". (1998, p. 450) Although "representation *of* similarity" may be sufficient to meet the requirements of role B, its inherent ambiguity with respect to the properties of the target (since here it is merely similarity which is being represented) makes it unsuited to role A. Both the utility and ambiguity of the "representation *of* similarity" will become clearer when we discuss representations posited in theories of perception in the following sections.

We had an intimation of the ambiguity inherent in second-order isomorphism earlier in considering the engineering example. Since the data structure representing the surface of the vehicle, is, at the time of writing, instantiated in the distribution of electric charges on RAM, orientation of magnetic domains on hard drive and ink on paper in figure 3d, the only representational content that can be supplied is that which is common to figure 3d, RAM and hard drive. The content is in this respect ambiguous. We can now see that this ambiguity, with respect to the properties of the target, arises inevitably from the fact that the data structure, that is the representation, is second-order isomorphic to its target. The representation is *of* similarity,

not *by* similarity, and there is no requirement that the representation and its target have any properties in common.

In summary, physical science and engineering provide examples of representations, instantiated as data structures in von Neumann computers, which are used in the prediction of the behaviour of complex systems. By virtue of being second-order isomorphic to their targets, such representations provide a basis for the "surrogate reasoning" on which the efficacy of the simulations depends. However, although promising candidates for role B, an inherent ambiguity with respect to the properties of their targets appears to render purely-second-order isomorphic representations unsuitable for role A. The question now, addressed in the following sections, is whether the representations posited by cognitive science are better candidates for role A.

For comparison with the following, it is useful to position the engineering example in terms of the distinctions drawn in Section 3.2.3. Here we have an example of a representation instantiated in a rule-based symbol-system, which is digital, time-step, static, local and objective. In this respect it aligns with the paradigm of a representation in GOFAI.

4.3 Representations in Marr's orthodox computational theory

Although Marr's theory (1982) has been subject to major revisions eg (Hayward & Tarr, 2005, p. 54), more recent computational theories of vision can be seen as building on this seminal work. His theory, in which vision is explained in terms of the processing of representations, is an exemplar of orthodox cognitive science and remains the subject of much discussion. Moreover, his 2½-D sketch is still cited by proponents of representationalist direct realism as the kind of representation that might be involved in such a theory.

Marr posits a sequence of three representations: primal sketch, 2½-D sketch and 3-D model representation. The first is constructed from the pattern of stimulation on the retina, the second is a transformation of the first and the third of the second. See figure 5 below. The last of the three

representations, the 3-D model representation, is designed for the purpose of object recognition.

However, the 2½-D sketch, which serves as an intermediary in the construction of the 3-D model

representation, appears to be the most likely candidate for role A, that is in Marr's terminology

an explanation of "pure perception" which is prior to "interpretation":

> … the 2½-D sketch provides a viewer-centered representation of the visible surfaces in which the results of the processes described in Chapter 3 [eg stereopsis, shading and motion perception] can be announced and combined. The construction of the 2½-D sketch is a pivotal point for the theory, marking the last step before a surface's interpretation and the end, perhaps, of pure perception. (Marr, 1982, p. 268)

He goes on to refer to the 2½-D sketch as:

> … an internal representation of objective physical reality that *preceded* the decomposition of the scene into "objects" and all the concomitant difficulties associated with object recognition. (p. 269)
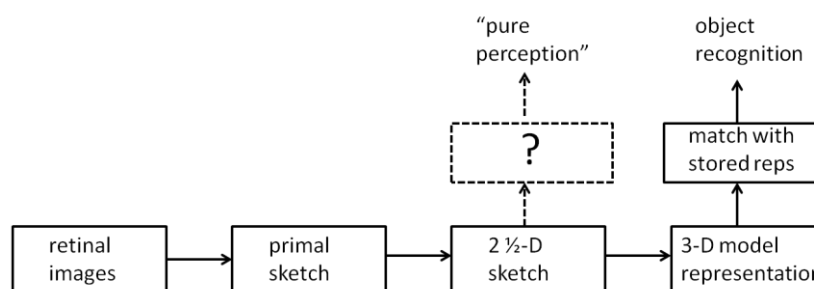
Figure 5

Tye, a leading proponent of Strong Representationalism, agrees that it is to the 2½-D sketch to

which representationalism might look to find a supplier of representational content:

> The representation we [as perceivers] share here has a content much like that of the 2½-D sketch posited by David Marr in his famous theory of vision (1982) to which further shape and color information has been appended (for details see Tye 1995)… Representationalists sometimes claim that it is here at this level of content that qualia are to be found. (Tye, 2015)

Tye's 1995 reference is to his *Ten Problems of Consciousness* in which he writes:

> The obvious conclusion is that there is a stage in *normal* vision in which representations of the local elements of the visual field are combined or grouped into an overall representation of the surfaces visible from the given point of view…. It seems plausible to suppose that it [the above representation] has the structure of a grouped array like Marr's 2½-D sketch (Marr 1982), whose cells are devoted to specific lines of sight relative to the viewer (with different cells devoted to different lines). (Tye, 1997, p. 210)[156]

Here in the 2½-D sketch we have a plausible candidate for role A, which we will examine in detail. However, it is important to note that Marr is primarily concerned with the object recognition problem, in our terminology representations for role B, rather than giving an account of perceptual experience. The difference is brought out by the solid and dotted paths in figure 5. Object recognition, the solid path, is explained in terms of a process in which the 3-D model representation of an object is matched "to a stored three dimensional representation with which other knowledge is already associated" (Marr, 1982, p. 326), and does not involve any reference to perceptual experience.[157] In Marr's theory, the 2½-D sketch is an intermediary in the production of the 3-D model representation from the retinal image and the primal sketch.

As might be expected, since Marr's 2½-D sketch is a data structure[158] which could be instantiated within a rule-based symbol-system, it bears a close resemblance to the representation in the engineering simulation running on a PC or mainframe. Indeed, in discussing the options for shape representations, Marr (1982, p. 274) identifies the same three design choices as would face the programmer in the engineering example: type of coordinate system, nature of the shape primitives, and the overall structure (hierarchical or flat) of the representation.

With respect to the first two points Marr decides that:

---

[156] See also Tye's Consciousness Color and Content (2000, p. 71)

[157] Marr's reference to perceptual experience is vague and it is unclear how experiences map onto the stages of processing posited by his theory. See section 7.2 of Marr (1982) for his brief discussion of this issue.

[158] As discussed below, Marr is explicit on this point.

> The coordinate system must be viewer centered, and the shape primitives must be two-dimensional and specify where the local pieces of surface are pointing. (p. 274)

By "viewer-centered" coordinate system, Marr means that the coordinates specifying the orientation and distance of each shape primitive must be defined with respect to the viewer. In this respect, that the origin and orientation of the coordinate systems are to be defined independently of the object being represented, the 2 ½-D sketch parallels the engineering example. However there is a difference in the surface elements, that is "shape primatives" or "surface primatives" in Marr's terminology. Whereas the surface elements in the engineering example were quadrilaterals specified by the coordinates of their vertices, in the 2½-D sketch they are "surface primatives of one (small) size" [159], which are specified in terms of distance from the viewer and their orientation with respect to the "direction straight ahead". (p. 277) Marr suggests, for the purposes of discussion, a particular version of the 2½-D sketch in which:

> Depth may be represented by a scalar quantity r, the distance from the viewer of a point on the surface. Surface discontinuities may be represented by oriented line elements. ….surface orientation may be represented as a vector (p,q) in two-dimensional space, which is equivalent to covering the image with needles [Marr's reference to needles is useful in interpreting the diagram below]. (1982, p. 277)

The viewer centered coordinate system, and a surface primitive which is defined in terms of its distance r and orientation (p, q) with respect to the viewer, are depicted in figure 6c. Figure 6b, adapted from Marr along with his original annotation, shows how the cylindrical object in figure 6a can be built up from these surface primitives. (The figures are labeled so as to correspond to figures 3a to 3d in Section 4.2)
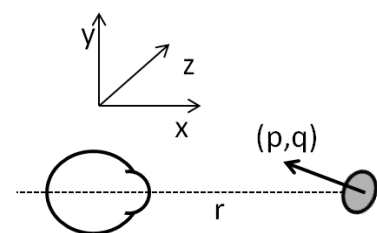
Figure 6 c

---

[159] Although Marr uses "generalized cones" as the volumetric primitives within the 3-D model representation (Marr, 1982, p. 303), the geometry of the surface primitives in his example of the 2 ½ -D sketch is not defined. He pictures such primitives as "fish scales" covering the surface. (p. 302)

Figure 6a

"[T]his symbolic representation is used to show the surface orientations of two cylindrical surfaces in front of a background orthogonal to the viewer. The full 2½-D sketch would include rough distances to the surfaces as well as their orientations; contours where surface orientations change sharply, which are shown dotted; and contours where depth is discontinuous (subjective contours) which are shown with full lines" (Marr, 1982, pp. 129 fig 3-12b)

Figure 6b

We now need to understand how the representations of the individual surface elements are brought together in the 2½-D sketch; this is the third of Marr's three design considerations. Just as in the case of the representation in the engineering example, the 2½-D sketch is a data structure. Marr is explicit on this point: "We have christened the resulting data structure the 2½-D sketch." (n.d., p. 44) He does not discuss the organization of the data structure in detail but does suggest that it should be flat, rather than hierarchical as in the 3-D model representation. (p. 302) Indeed since, as in the engineering example, each surface element has 8 neighbours, a flat, non-hierarchical data structure, in which the data triples (r, p, q) corresponding to spatially adjacent surface elements are logically adjacent in the data structure, is the simplest option.

Marr does not give an example of the 2½-D sketch in data structure form. However, since we need to understand what sort of thing the 2½-D sketch is, how it represents, and what it represents, it is useful to construct a data structure corresponding to figure 3d in the engineering example. Marr notes that the "full 2½-D sketch would include rough distances to the surfaces as well as their orientations", along with contours where surface orientations change sharply and where depth is discontinuous. (1982, p. 129) Figure 6d is a data structure which includes

distances and orientations but not contours of discontinuity. Here, in ink on paper, we have an instantiation of a 2½-D sketch, albeit somewhat simplified.

91	999	0	91	0	560	91	0	560	91	0	560	91	0	560	91	0	560	91	999	0
92	999	0	70	89	18	64	35	18	62	0	18	64	-	18	70	89	18	92	999	0
93	999	0	71	89	18	65	35	18	63	0	18	65	35	18	71	89	18	93	999	0
93	999	0	94	0	560	72	89	18	66	35	18	64	0	18	66	35	18	72	89	18	94	0	560	93	999	0
95	999	0	69	113	18	60	58	18	56	26	18	55	0	18	56	26	18	60	58	18	69	113	18	95	999	0
96	999	0	70	113	18	61	58	18	57	26	18	56	0	18	57	26	18	61	58	18	70	113	18	96	999	0
97	999	0	71	113	18	62	58	18	58	26	18	57	0	18	58	26	18	62	58	18	71	113	18	97	999	0
72	113	18	63	58	18	59	26	18	58	0	18	59	26	18	63	58	18	72	113	18

Figure 6d

Although the numerical values in figure 6d do correspond, on a plausible interpretation of Marr's proposal, to the distance and orientation of the surface primitives in figure 6b, they are unimportant. The crucial point is that if we are to take seriously the notion of a data structure, that is a representation, which represents the layout of surfaces in terms of their distance and orientation with respect to the viewer, then the "r, (p,q) scheme" or something of this sort is required. Put another way, the sort of representation we are talking about is something like figure 6d, that is a structured arrangement of tokens which can be assigned values which in some way correspond to distance and orientation of surfaces.

We can now see that the 2½-D sketch represents the cylindrical object by virtue of second-order isomorphism. Consider the three surface primitives SP1, SP2 and SP3 and the corresponding section of the data structure, that is the three triples (r_1,p_1,q_1) (r_2,p_2,q_2) and (r_3,p_3,q_3), shown in figure 7.

r_1	p_1	q_1
72	89	18

r_2	p_2	q_2
94	0	560

r_3	p_3	q_3
93	999	0

Figure 7

Clearly there is no similarity, concrete or abstract, between a surface primitive and the corresponding triple. However there is a one to one correspondence,

that is isomorphism, between the surface primatives and the triples, such that result of an operations on a surface primitive, that is measuring its distance from the viewer with a tape, corresponds to the results of analogous operations on the corresponding triple, that is applying a particular computation. In this case, the particular computation on the three triples would give the results 72, 94 and 93, which preserve the proportions of the corresponding distance measurement on the surface primitives; in this case the computation is trivial, simply the extraction of the first element, ie r, from the triple. For example the distance measurements on SP1 and SP2 would be in the proportion 36:47, and the results of the computations on the corresponding triples, 72 and 94, are in the same proportion.

There is also a proportion-preserving correspondence between the surface primatives and triples with respect to orientation, but here the computation on the triples is more complex. The actual computation used to produce figures 6d and 7 is a speculative interpretation of Marr's suggestion and the numerical values are unimportant, but the fact that the computation is in this case non-trivial should be noted.[160]

The key point, paralleling that made in the engineering example, is that figure 6d is a representation of the cylindrical object in figure 6a. It is not just an illustration of such a representation, it is the real thing.[161] Figure 6d is a 2½-D sketch (albeit simplified) instantiated in

---

[160] There are a number of options, but it is clear that Marr is taking the vector (p,q) to be understood as a point in "gradient space" (p. 277) a notion which he discusses earlier in the text and describes as "an elegant way of representing surface orientation". (p. 242). Figure 6d and 7 have been constructed on this basis. Taking a plane normal to the direction of view as the datum, p corresponds to the degree of rotation of the local tangent plane about a vertical axis, and q similarly with respect to rotation about a horizontal axis. Strictly, p and q are the tangents of their respective angles, but in the data structure the values have been multiplied by 100, rounded and limited to 999 (the tangent of 90 degrees is infinite). In this way, the three values of p (89, 0, 999) correspond to surface elements "facing right" at an angle of 42°, 0° and 90°. Similarly the three values of q (18, 560, 0) correspond to surface elements "facing up" at -10°, 80°, and 0°. For simplicity and space all values in the table are positive, but +ve and –ve values are needed to cope with left/right and up/down facing elements.

[161] It might be objected that figure 6d, a table of numbers which looks suspiciously like an Excel spreadsheet, cannot really be an instantiation of a 2½-D sketch. However, although Marr recognizes that numerical variables and coordinates may be difficult to accept in the context of neuron firings (1982, p. 80), he talks explicitly in terms of "tokens" which "can be assigned values of attributes like orientation... and position" (p. 44).

ink on paper, which is at the time of writing also instantiated in electric charge on RAM and magnetic domains on hard drive. Marr posits such representations instantiated in activation levels on neurons in the brain as the basis for an explanation of the efficacy of visual perception. Crucially, though different in physical realization and the conventions used, all such instantiations provide the same representational content.

Against the charge of labouring a point, our approach serves as an essential Dennett-style "intuition pump", which must be worked hard if we are to understand what the representations posited within cognitive science **actually are, what they represent, and how they represent**. Much of the discussion defending representationalist direct realism proceeds on the basis of the notion of representation and representations in the abstract, or in general terms, and needs to be grounded on a detailed examination of concrete examples thereof.

Here in Marr, as in the engineering simulation, is a case of surrogate reasoning on the basis of representations which are second-order isomorphic to their targets.[162] We have representation *of* similarity rather than representation *by* similarity; the representation and its target do not have any properties in common, be they abstract or concrete. Indeed the differences between the two representations of the layout of surfaces are differences in detail rather than principle. Consequently it appears that the 2½-D sketch is no better suited to role A, than was the representation in the engineering example.

---

[162] To see how isomorphism, as opposed to mere law-like relations or causal co-variation, make information available in usable form, consider the example of the square and trapezoidal surface elements in Section 4.2. Suppose [200, 250, 900, 124] was an arbitrary code for a square of particular area, position and orientation, and [250, 200, 876, 392] for a particular trapezium. Although, in principle, the surface structure of the vehicle could be encoded in this way and the simulation software modified accordingly, the scheme would be completely unworkable in practice. The problem is that although all the necessary information is available it is not in a readily usable form; the storage and processing requirements would be prohibitive as a massive look-up table would be required. By using representations that are isomorphic to their targets the information becomes extractable. Kulvicki discusses the notion of "extractability" in the context of perceptual representations and argues that although "isomorphism plays no role in fixing the contents of perceptual states, it is essential to those states performing their function, which is making information about the ambient environment available to the cognitive faculties of the individual." (Kulvicki, 2004, p. 381)

This is not a criticism of Marr. He is not setting out to account for perceptual experience and is certainly not proposing the 2½-D sketch as a candidate for role A. Rather he is concerned with role B, that is explaining perceptually guided behaviour.[163] However our conclusions with respect to role A are at odds with those who posit the 2½-D sketch as a representation within a representationalist version of direct realism.

The 2½-D sketch is difficult to position in terms of the distinctions drawn in Section 3.2.3: digital v analogue, time step v continuous time, local v distributed; though as discussed in the next section, it is again objective[164] rather than action orientated. The difficulty arises from the different levels at which an information processing system can be understood. Here Marr famously identified three levels:

> At one extreme, the top level, is the abstract computational theory of the device, in which the performance of the device is characterised as a mapping from one kind of information to another ….In the centre is the choice of the representation for the input and the output and the algorithm to be used to transform one into the other. And at the other extreme are the details of how the algorithm and representation are realised physically … (1982, p. 24)

Whereas the representation in the engineering example, say it is programmed in C++ and running on a PC, is specified at all three levels, the physical realisation of the 2½-D sketch is not specified. Moreover although, as we have seen, the 2½-D sketch is similar to the engineering example at the top two levels, there is no stipulation as to how it is realised physically. Although Marr talks in terms of algorithms, this should not be taken to imply that he supposes that the 2½-

---

[163] The 2 ½-D sketch serves as an intermediary in the production of the 3-D model representation which enables the organism to "recognize an object by its shape" and thus "evaluate its significance **for action**. [emphasis added]" (1982, p. 326)

[164] Marr describes the 2½- D sketch as "viewer-centered", in contrast to the "object-centered" 3-D model representation, however the distinction is only with respect to the coordinate systems used. Both representations are objective as opposed to "action-oriented" in the sense that they represent objective features of physical world, as opposed to a particular animal's opportunities for action. (1982, p. 269) Note, lest there appear to be an inconsistency with the above, that the 3-D model does not represent opportunities for action, rather these need to be evaluated from the 3-D model in a further stage of processing.

D sketch is instantiated in a rule-based symbol-system. Indeed he discusses, for example, the implementation of Gaussian filters in physiological terms. (1982, p. 64)

We will draw further on Marr's three levels in the following sections. For now, we suggest, for later consolidation, that the requirements for role B can be specified at the top level in terms of inputs and outputs, whereas role A puts requirements on the second level in terms of the relationship between representation and target. It seems that neither role places a requirement on the bottom level of physical realisation.

The question is now whether recent development in cognitive science, particularly connectionist models of the perceptual process, can provide a more suitable candidate for role A.

4.4 Representations in Churchland's connectionist theory

Whereas Marr stands out as the seminal figure in terms of orthodox computational theories of vision, there is no obvious champion on the connectionist side. However the Churchlands have written extensively on representation in connectionist networks and their work, particularly Patricia Churchland's *Brain-Wise* (2002) and Paul Churchland's *Plato's Came*ra (2012), provides the basis of our discussion. Churchland proceeds on the assumption, which she recognizes as such[165], that perception involves representations and computations thereon:

> However brains work, much of what they do involves *representing* – representing the brain's body, features of the world, and some events in the brain itself. Performing computational operations on those representations serves to extract relevant information, make decisions, remember, and move appropriately. (2002, p. 273)

However, Churchland would criticize the orthodox approach, in terms of algorithms and data structures in rule-based symbol-systems, as taking too little account of the findings of

---

[165] Churchland recognizes that these are "working assumptions" and "it is too early in the game to be very confident that "*representing* reality" is the right way to describe the workings of the mind-brain." (p. 270)

neuroscience.[166] In contrast she argues for a brain-friendly approach to computation and representation in terms of the more biologically plausible connectionist networks.

Churchland (2002, p. 293) poses the key question "How can neuronal activity represent something?" and offers an explanation in terms of "the vector/parameter space approach to representation", which she illustrates using a three stage artificial neural network, developed by Cottrell et al, that can perform recognition tasks on photographs of human faces. See figure 8[167]. As Churchland points out, although the brain is massively complex and very little is known about how neurons do in fact represent faces, the simple Cottrell network is useful for demonstrating the principles that might be involved. Moreover, as discussed in Section 4.5, there is good reason to suppose that these principles apply to the more complex networks which are now at the cutting edge of connectionist AI.
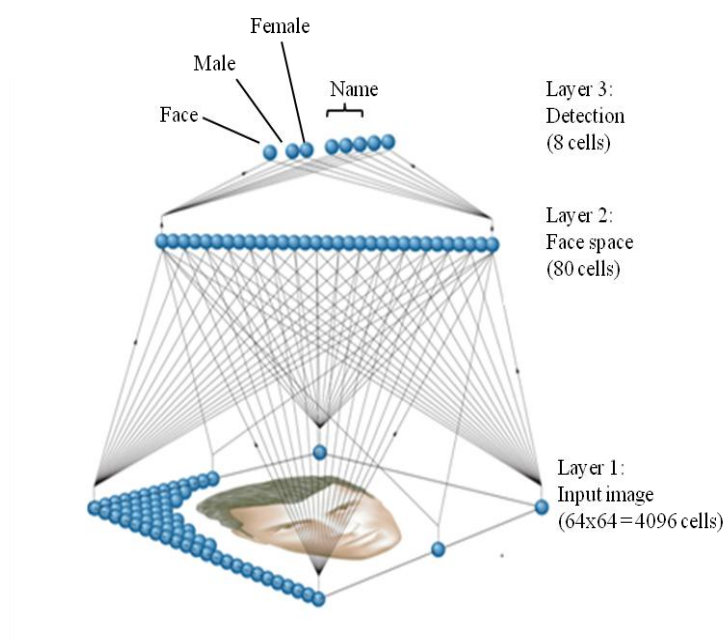
Figure 8

---

[166] Though Marr does make considerable reference to neuroscience and is some way removed from the paradigm of the "brain-averse" approach with which Churchland takes issue. In this context, she makes particular reference to Fodor and Pylyshyn. (Churchland, 2002, p. 273)

[167] This figure is based on (Churchland & Churchland, 2007) but with changes: original annotated in Italian.

The Cottrell network is structured as follows:

> Face net's input layer (for our purposes, a pretend retina) is a (64 × 64)-pixel grid whose elements each admit of 256 levels of activation or "brightness" according to the light reflected from the region in the photo to which it is sensitive. The network's input consists of grey scaled photographs. …. Each input unit projects a radiating set of axonal end branches to each and every one of the 80 units in the second layer, and this layer maps an abstract space of 80 dimensions (a dimension for each unit) in which the input faces are explicitly coded. …. The second layer projects to an output layer of merely eight units. These output units have their connection strengths carefully adjusted so that the units can make a number of discriminations: first discriminating between faces and nonfaces; second, discriminating between male and female faces; and the third, responding with the person's "name" (actually an arbitrarily assigned binary code) when presented with any face that the network "got to know" during training. (Churchland, 2002, p. 294)

Since we are only concerned with how, in principle, such networks represent, we do not need to consider the training process by which the connection weights are adjusted. (Though, it should be noted, that training strategies and algorithms are very important topics in connectionist research and development. (Hassoun, 1995) (Heaton, 2015)) We will assume that the network has been trained, and the weights have been adjusted, such that the network can achieve a good level of success in terms of putting names to faces when presented with novel photographs of the faces in the training set.

Before examining the representations in such a network, we need to clarify a "crucial ambiguity". Churchland (2002, p. 296) notes that the term "representation" is sometimes used to refer to "cognitive events happening *now*, such as a visual perception", and other times to refer to "the *capacity* (*not* now exercised) to have appropriate cognitive events". Representations in the first sense are transitory "patterns of activity" say neuron firings, and in the second sense more long lasting "configurations of connection weights". We are concerned with representation in the first sense, that is with representations as providers of the representational content

associated with particular perceptual experiences, rather than in some way instantiating the background conceptual framework which makes such perceptions possible.[168]

In discussing representation in Cottrell's network, Churchland presents "the basic story" thus:

> The values in the input vector reflect the gray-level values in the photographs, and the configuration of connection weights in the middle-layer vectors embodies what is task-relevant in various aggregations of input values. Input vectors are pushed through the configuration of weights, transformed into abstract representations in a high dimensional "facial parameter space." These vectors are in turn pushed through the last layer of weights with the resulting output vector representing answers to "Is it a face or not?" "Is it a male or female?" and "Who is it?" (2002, p. 299)

Here, as in Marr's proposal, we have a sequence of three representations. Indeed the three layers of Cottrell's network are, in terms of function, suggestive of Marr's primal sketch, 2½-D sketch and 3-D model representation respectively, though there are fundamental differences in the nature of the computations involved and the fit is very far from exact. The third layer is essentially a set of what we might call "object recognizers", variants on the supposed "Grandmother neuron"[169], which fires in recognition of male faces, female faces, Bob's face etc. Marr's "3-D model representation" (1982), along with the associated algorithms and library of previously stored representations, performs the same object recognition function, and could be configured to provide a similar set of recognizers each of which provides a single bit of information, that is male/not male, Bob/not Bob etc. However, the 3-D model representation represents objects in a much stronger sense[170] than does the third layer in the connectionist network which merely indicates the presence of particular objects. Indeed Ramsey (2007, p. 132), taking issue with Dretske, classifies such third layer representations as "receptors" or

---

[168] It is representations in the second sense that constitute the "Landscape of Abstract Universals" that feature in the sub-title of *Plato's Camera* (Churchland P. M., 2012)

[169] The supposed "grandmother neuron" which consistently and reliably fires when Grandmother is present in the visual field, after being rejected as laughably crude, has been given a new lease of life as the Jennifer Aniston cell. See *Invariant visual representation by single neurons in the human brain* (Quian Quiroga, et al., 2005)

[170] The 3-D model representation is an information rich object-centred representation of the layout of the surfaces of the object in terms of "generalized cones" (Marr, 1982, p. 318)

"indicators" and argues that they are not true representations. Whether or not the single-bit object

recognizers at the third layer should be regarded as true representations is moot, but they are

clearly incapable of supplying the representational content necessary to account for the rich

phenomenal character of perceptual experience such as that of seeing Bob. To find a possible

candidate for this role, role A, we need to turn to the lower layers of the network.

It is at layer two that we find the distributed representations that are usually taken to be

characteristic of connectionist networks. Churchland explains distributed representations as

follows:

> … some values are coded by a population of neurons whose members are active to
> different degrees across a range of properties (vector coding) ….[for example] the brain
> could represent the face of Woody Allen with a particular *pattern* of responses in the
> population, and the very same population of neurons, but with a different pattern of
> responses, may represent the face of Ghandi and the face of Castro. (2002, p. 289)

By contrast, in the case of local representation, "the face of Woody Allen may be coded by a

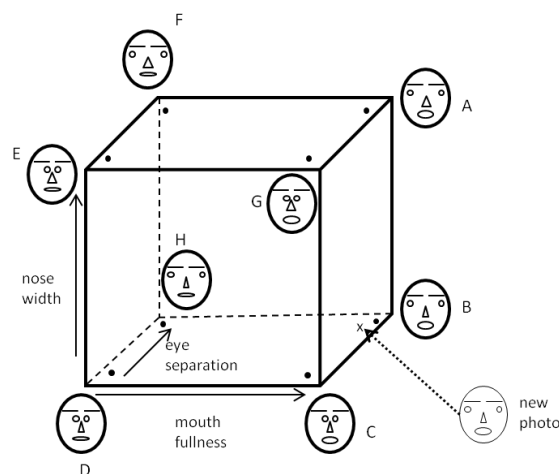single neuron, and this neuron only fires when and only when a Woody Allen face is presented."

(p. 289) The supposed "Grandmother neuron" and the object recognizers at layer three of

Cottrell's network are local in this sense.

Churchland explains the distributed representations at layer two using three dimensional

diagrams of the 80 dimensional state space; the other 77 dimensions have to be left to the

imagination. Such diagrams:

> … illustrate the idea that faces vary along a number of dimensions, represented as axes of
> the state space, and that the system might code for faces using vectors whose elements
> might represent such features as distance between the eyes, fullness of the mouth, and
> width of the nose. (2002, p. 292)

Figure 9, adapted from Churchland (2002, p. 292), relates to a simplified version of Cottrell's

network with three, rather than eighty, cells at layer two. The three cells are assumed to be

activated at a level determined by mouth
fullness (m), eye separation (e) and nose
width (n).[171] Thus any face presented to the
network will activate the three cells at levels
determined by the presence of those three
features, and the state of the system can be
defined by a three element vector (m,n,e)
which can be interpreted as a point in 3-D "face space".

Figure 9

The utility of such a network for the purpose of object recognition is that similar faces are
represented by nearby points in face space, and dissimilar faces by distant points. Consequently,
if a previously unseen photograph is presented to the network, the face can be recognized by
determining its nearest neighbor in face space. Suppose a photograph showing a face with a full
mouth and widely separated eyes but narrow nose is presented to the network. The face in this
photograph would be represented by point X in face space, and recognized as a photograph of
face B by virtue of its proximity in face-space. In the case of the complete Cottrell network the
explanation would be in terms of an 80 dimensional face space in which the dimensions
correspond to abstract, and not intuitively significant, properties of the faces (2002, p. 298).
However, the correspondence between similarity of faces and the proximity of points in face
space carries through to the more complex case.

Although the notion of an abstract multidimensional space is a useful tool for thinking about the
representations at layer two, such explanations must not be taken to suggest that connectionist

---

[171] This is a simplification, the features are not humanly intuitively recognizable. Talk of nose length etc is purely
illustrative. "Notice that each cell comprehends the *entire surface* of the input layer, rather than an isolated facial
feature, such as the nose." (Churchland, 2002, p. 298)

networks thereby have representational capabilities not found in the data structures within rule-based symbol-systems. Indeed Marr's 2½-D sketch could be described in terms of multidimensional abstract spaces[172], and the Cottrell network described in terms of data structures. Although Churchland talks in terms of "vectors" and "processing" (2002, p. 296), the distributed representation at layer two of Cottrell's network is, no less than Marr's 2½-D sketch, a data structure on which computations are performed, albeit weight-based rather than rule-based computations. [173]

It is clear that the distributed representations at layer two are an essential step towards object recognition at layer three of Cottrell's network. Indeed, as discussed below, recognition by adjacency in face space appears to be a particularly neat solution to the problem, and we will accept that such representations can in principle meet the requirements of role B. However with respect to role A, the layer two representations appear to be no better suited than the representations in the engineering example or Marr's 2½-D sketch.

There is no similarity, concrete or abstract, between a face and a pattern of cell activation levels. Rather, there is a similarity between the relationship between faces and the relationship between the corresponding patterns of cell activation levels. Specifically, there is a one to one correspondence between faces and points in abstract face-space, such that an operation on a face P, say an assessment of its likeness to face Q, corresponds to an analogous operation on the corresponding point p in face-space, that is computing a distance between p and q in face-space. The correspondence is such that certain relationships between the results of likeness assessments on the faces are preserved in the results of computations on the points in face-space. For

---

[172] See Palmer (1978, p. 283), for a discussion of "Multidimensional Spaces" and the generality of the notion.

[173] "The fundamental processing format consists of mere vector-to-vector transformations determined by the configuration of connection weights." (Churchland, 2002, p. 296)

example, consider three faces P, Q and R, and the corresponding points in face-space p, q and r. If face P is assessed to be more like Q than R, then point p will be closer to q than r. We have again a case of representation by second-order isomorphism and inference by surrogate reasoning. Indeed, here, in distinction to the previous examples which might be taken to have some features of first-order abstract isomorphism, the representations are purely second-order isomorphic to their targets.[174]

Here again, as in the engineering example and Marr's 2½-D sketch, it appears that although the distributed representation at layer two of the Cottrell network is a potential candidate for role B it cannot meet the requirements of role A: it cannot supply the representational content required of a representation in a direct realist theory of perception. To consolidate this point, we need to understand more clearly both the weakness of layer two representations as candidates for role A and their particular strengths with respect to role B. First we consider role A.

Figure 10a is again, as was figure 9, based on Churchland's original "Diagram of a face space" (2002, p. 292), but now with two changes. Firstly the targets in physical space and their representations in face space are shown separately, and their separation is emphasized by the dividing line. Secondly the axes of the three dimensions in face space have been re-labelled "cell x activation level" etc, rather than "mouth fullness", etc. As we will see, for our present purposes, the original labels would have been dangerously misleading with regard to the content of the representation.

The crucial point is again that there is no similarity, abstract or concrete, between the points in face space and the corresponding faces in physical space; there is no similarity between a and A

---

[174] As we will see in Section 4.5, Edelman (1998, p. 449) might regard the 2½-D sketch as a "set of measurements that collectively encode the geometry and other visual qualities" of an object, and thereby first-order abstract isomorphic to its target.

or b and B etc. See figure10a. Moreover the points in face space do not provide any information about the properties of their targets. The information available in face space, that is encoded in the activation levels of the cells at layer two, is in terms of the similarities and differences between faces A , B, C etc with respect to three unspecified properties; note "unspecified". This is why it would have been misleading to retain the original labels on the axes: it would lead us to overstate the information that is available within the representation. Labelling the axes "mouth fullness" etc, that is associating the firing rate of cell x with a particular property of faces, would require information that is only available outside the representation and the network within which the representation is instantiated. Indeed it (labelling the axes) would require a designer's-eye view of the target, network and representation.
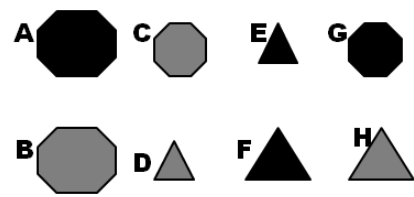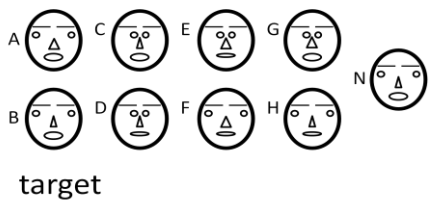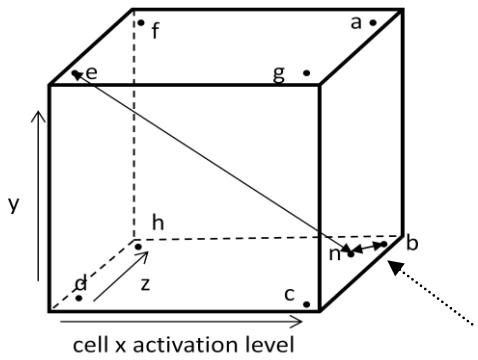

target

representation

Figure 10b

$$sqrt((x_n-x_b)^2+((y_n-y_b)^2+((z_n-z_b)^2)$$

Figure 10a

(It is important to note the fundamental distinction between this case, where the nature of the perceptual process is the point at issue, and the engineering case where we do have a designer's

eye view. [175] However, given the strength of our pre-critical intuitions, it is all too easy to overlook this crucial distinction and to take ourselves, the perceivers in this case, to be in a similar position to the designer of the simulation model in our discussion of the engineering case who has the same epistemic access to both the representation and its target. Indeed, it is useful to be reminded of Strawson's warning, see Section 1.3, regarding the grip of "common-sense non-representational realism" even on those "who are intellectually convinced of its falsity". (1979))

The ambiguity with respect to the properties of the faces, concerning which information about relations is encoded in the second-order isomorphic representation at layer two, can be seen by considering the shapes shown in figure 10b. Although the faces A, B, C etc do not have any properties in common with the shapes **A**, **B**, **C** etc, they do share the same similarity and difference relationships. For example A is similar to C in terms of two properties and similar to G in terms of one property, and the same relationships hold between **A**, **C** and **G**. (The properties mouth fullness, nose width and eye separation correspond to sidedness, shading and size respectively.) Since second-order isomorphic representations are representation *of* similarity rather than *by* similarity, and the faces and the abstract shapes have the same similarity and difference relations, the parameter space is in figure 10a is no less a representation of the abstract shapes than it is a representation of the faces. Here we can see why the representation cannot supply representational content which represents faces as having the properties that they (faces) actually have. The content is limited to that which is shared by the faces and the abstract shapes, and they have no properties in common.

As O'Brien and Opie put it in a paper arguing for a "structuralist theory of mental representations" and in the context of a discussion of the Cottrell network":

---

[175] This key point applies equally to our discussion of Marr (1982) in Section 4.3 and Edelman (1998) in Section 4.5. However, it is most clearly made here in the context of figure 10 and the designer's eye view.

As already stressed, resemblance of this kind [second-order resemblance] is independent of first-order resemblance, in the sense that two systems can resemble each other at second-order **without sharing properties**. [emphasis added]. (2004, p. 8)

We now turn to role B. At this point second-order isomorphic representations, merely representations *of* rather than *by* similarity, might appear to provide only a very tenuous grasp on the world and it is important to emphasize their utility for the purpose of perceptually guided behaviour. As we have seen, such representations are suited to the task of face recognition in the Cottrell network. Similar faces are represented by points that are close together in face space. Thus, in figure 10a, the similarity of faces N and B can be deduced from the proximity of points n and b in face space. Moreover "proximity" can be given a precise meaning in terms of the Euclidean distance between the points n and b, that is $sqrt((x_n-x_b)^2+((y_n-y_b)^2+((z_n-z_b)^2)$. So reasoning in terms of Euclidean distances in face space can be used as the basis of inferences about the similarity of faces in physical space. (The notion of Euclidean distance extends to any number of dimensions in an abstract space.) Here is surrogate reasoning at work.

Even these most basic capabilities could provide an organism with important perceptually guided behaviours. Supposing type B faces are friends and type E faces are foes, then on encountering a new face N, the appropriate behaviours could be initiated depending on the position of the corresponding point n in face space. If $((x_n-x_e)^2+((y_n-y_e)^2+((z_n-z_e)^2) < ((x_n-x_b)^2+((y_n-y_b)^2+((z_n-z_b)^2)$, then n is closer to e than it is to b in face space, and N is more like E (foe) than B (friend); thus avoiding action may be appropriate. It needs to be stressed that this piece of surrogate reasoning does not require, indeed would not benefit from, the association of x with mouth fullness etc. For the purposes of taking appropriate avoiding action the fact that the information required to, as it were, label the axes, is not available, is of no consequence.

Here we are only scratching the surface with respect to the perceptually guided behaviours that could be made on the basis of such representations. Even a simple Cottrell network would allow an organism to discriminate, with some degree of success, between male and female faces and initiate appropriate behaviours. Indeed a plausible evolutionary explanation of the development of such networks in the natural world can be given on the basis of the selection for such efficacious behaviours, building up from the simplest organisms to the human with a network of billions of cells. As Churchland puts it with respect to the explanatory power of representations in parameter space:

> What is so tremendously useful about *spatiality* here is that the space (3-D, 10-D or n-D) has a metric meaning that positions in the space can be specified as near each other or far from each other or in-between. And spaces admit of regions, volumes, paths and mappings. All of this makes it easier to conceptualize representations, relations between representations, and relations between representations and the world. (2002, p. 291)[176]

We will grant that such second-order isomorphic representations can, in principle, meet the requirements of role B, that is provide the basis for an explanation of perceptually guided behaviour. Importantly, our arguments with respect to role A do not require us to challenge the claims of cognitive science with respect to the efficacy of such representations in role B. However, neither need we endorse these claims.

Up to this point we have focussed on the particular requirements of role A, that is to supply the representational content needed to meet the requirements of the representationalist version of direct realism. However, as we noted in Section 4.1, there is a more general requirement on any representation-friendly theory of perception, direct or indirect, with respect to the richness of experience. The representation must be able to supply sufficient representational content to

---

[176] In his recent monograph, *Plato's Camera – How the Physical Brain Captures a Landscape of Abstract Universals,* Paul Churchland (2012) builds on the notion of "mappings" within thousands of complex multi-dimension parameter spaces, corresponding to the synaptic weights and activation levels of cells within the human brain, to propose a comprehensive explanation of all our "cognitively informed behaviours". (p. 4) As the title might suggest, he is making a further claim with respect to the explanatory power and scope of such representations to

account for the richness of the phenomenal character of our perceptual experience. Now the requirement is quantitative rather than qualitative: it concerns the amount of information (bits) available, rather than what the information is about. Here again there is a difference in the requirements on the two roles.

To address this point we need to consider the information available at each of the three layers of Cottrell's network. See figure 8. Although it is treated merely as the input stage and given little attention in Churchland's discussion, the "64 × 64 pixel grid" of cells at layer one, which she calls a "pretend retina", is a representation in its own right. Moreover, in terms of accounting for the richness of perceptual experience, it appears to be better suited than representation at layer two by virtue of its greater information content. If as Churchland suggests, there are 256 levels of activation, then the information content available at the 4096 cell layer one is $4096 \times 8 = 32,768$ bits. (The factor of 8 arises since 8 bits of information are needed to distinguish between 256 states: $2^8 = 256$.) On the other hand, the information available at the 80 cell layer two would be proportionately less at only $80 \times 8 = 640$ bits. And at layer three, where there are 8 cells and only two (significant) activation levels the information available is at most $8 \times 1 = 8$ bits[177] corresponding to the presence of particular faces.

As discussed further in the next section, this successive reduction in the bit-count through layer 1, 2 and 3 is not just an accidental feature of the Cottrell network, but essential to the face recognition process. In these terms, we can understand the network as a mechanism for separating the information needed for classification and appropriate action at layer three from the mass of pixel-level information at layer one. These considerations with respect to information content clearly suggest that the requirements of roles A and B differ, not only in the type of

---

[177] This is a simplification since not all combinations are possible. A face cannot be both, say, Bob and Mary.

information that is to be delivered, but also in the quantity thereof. Indeed, such considerations raise the question as to why it should be supposed that the layer two representations are any better suited to role A than are the unprocessed pixel level representations at layer one; the same question arises in terms of the 2½-D sketch and retinal image in Marr's theory.

Before moving on to address these issues and consolidate our conclusions with respect to the requirement on representations in the two roles, it is useful to classify the representations at layer two of the Cottrell network in terms of the distinctions drawn in Section 3.2.3. The layer two representations are analogue rather than digital, time step rather than continuous time, static rather than dynamic[178] and distributed rather than local. The distinction between objective and action-oriented representations is not always clear cut; the one type tends to merge into the other. Although paradigmatic examples of objective representations could have been given (say a network which recognizes abstract shapes), our example of recognizing friendly and unfriendly faces could be regarded as involving action-oriented representations.

4.5 Direct realism and second-order isomorphic representations

Starting on the solid ground of the representations used in engineering simulations, and moving on to examples taken from orthodox and connectionist cognitive science, we have concluded that these representations are second-order isomorphic to their targets and could, as such, provide the basis for an explanation of perceptually guided behaviour. This much is relatively uncontroversial. However our claim that such representations cannot supply the representational content required of a representation in a direct realist theory of perception needs further development

---

[178] Time step and static, since the Cottrell network is a feed-forward network which settles into a stable state of activation levels depending only on inputs at Layer 1.

Edelman's target paper *Representation is representation of similarities* (1998) is particularly

useful at this point in our discussion, since he addresses second-order isomorphism and its

"philosophical implications" in some depth; he has also written two monographs on

representations in a theory of vision (1999) (2008). In the target paper Edelman presents for

discussion a theory of visual representation and object recognition in which:

> … a shape is represented internally by the responses of a small number of tuned modules, each broadly selective for some reference shape, whose similarity to the stimulus it measures. This amounts to embedding the stimulus in a low-dimensional proximal shape space spanned by the outputs of the active modules. This shape space supports representations of distal shape similarities that are veridical as Shepard's (1968) second-order isomorphisms (i.e., correspondence between distal and proximal similarities among shapes, rather than between distal shapes and their proximal representations.) (1998, p. 449)

Although there are significant differences between the system that Edelman proposes, and the

Cottrell network discussed by Churchland, they are similar with respect to the nature of the

representations involved. The "tuned modules", are the equivalents of the layer two cells in

Cottrell's network, and the "low-dimensional proximal shape space" the equivalent of

Churchland's 80 dimension "face space". In terms of the distinctions drawn in Section 3.2.3,

such representations in "proximal shape space" are, again, analogue rather than digital, time step

rather than continuous time, static rather than dynamic and distributed rather than local. Where

they stand between paradigmatic objective and action-oriented representations is less clear. As

we see below, they are representations for the purpose of "clustering by natural kinds".

In our terminology, the main thrust of Edelman's paper is to argue that such second-order

isomorphic representations are not only sufficient, but also necessary, to meet the requirements

of role B, and his concern with role A is peripheral. However his arguments and conclusions are

significant with respect to the latter. Figure 11, taken from Edelman and retaining his annotation,

serves to illustrate by virtue of its stark simplicity how second-order isomorphic representations which might meet the requirements of role B, are clearly unsuited to role A.
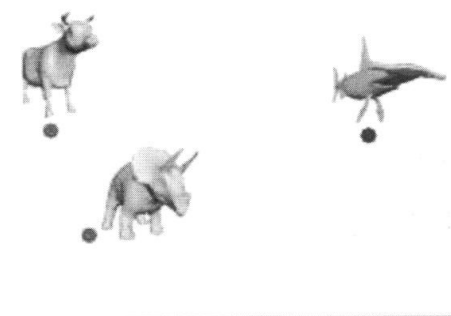
[The figure shows:]Clustering by natural kinds, and a representation of it that fulfils the requirement of second-order isomorphism, according to Shepard (1968). The deposition of the tokens corresponding to the three shapes in this illustration in the proximal representation space (bottom) reflects the disposition of the shapes in the distal shape space (top); the shapes of the tokens are irrelevant to their representational capacities. (Edelman, 1998, p. 450)
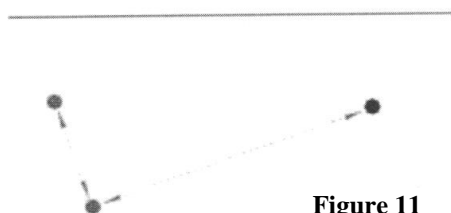
Figure 11

The crucial point is that although the disposition of the tokens in "proximal representation space" may serve as the basis for the categorisation and recognition of the corresponding shapes in "distal shape space", role B, such representations fail to meet the requirements of role A on two counts. Firstly, although there is, arguably, sufficient information (bit count) in the distribution of the tokens in shape space to account for object recognition, there is clearly insufficient information to account for the richness of perceptual experience. Indeed, Edelman recognizes that information reduction is an essential part of the object recognition process. He talks in terms of "dimensional reduction" (1998, p. 453) and the problem of separating "the chaff (pixel level information) from the wheat (classification information)." (p. 462) However although the chaff at the pixel level may be unnecessary, and indeed obfuscatory, with respect to role B, it would appear to be necessary to account for the richness of perceptual experience.

Secondly, with respect to the particular requirement of role A, such representations cannot supply representational content which represents objects as having the properties that they (the objects) actually have. In second-order isomorphism, that is representation *of* similarity, "only

certain relations between the objects – not the shapes of the individual objects themselves– need be represented." (Edelman, 1999, p. 27) Moreover Edelman argues that, for the purposes of role B, representation *by* similarity is not only unnecessary but also counterproductive: it amounts to "a reconstruction of the visual world" which "only attempts the absurdity of putting off until later the whole process of pattern recognition". (1998, p. 450) It is not that second-order isomorphic representation is an inferior form of representation which suffices for role B but not Role A. Rather, along with information reduction, representation *of* (rather than *by*) similarity is a particular requirement of role B which is not shared by role A. However there is an asymmetry between the two requirements. Once "the chaff" of pixel level information which might be taken to be necessary for role A, has been lost in the production of a second-order representation for the purposes of role B, it cannot be regained; the loss of information is irreversible. (Importantly, we make no positive claims as to the nature of representation, if any, which would meet the requirements of role A. See Section 1.2.4 where we distinguish between positive and negative claims and avow our modesty with respect to the latter.)

These considerations, particularly Edelman's references to "pattern recognition" and "dimensional reduction", help us to understand more clearly the differences between the two roles. In role B the requirement on the representations is to provide an explanation of perceptually guided behaviour in terms of pattern recognition, which in turn can be understood as a process of dimensional reduction from the sensory input to appropriate behaviors. In Churchland's example (2002, p. 294), 4096 dimensions at layer one ("a pretend retina") are reduced to 80 at layer two and 8 at layer three. The 8 dimensional output at layer three could then provide the basis for initiating perceptually guided behaviours appropriate to the presence of, say, male or female faces. Marr's theory could be presented in similar terms. That is successive dimensional reductions through the primal sketch and 2½-D sketch to the 3-D model

representation in which objects are represented as a number of "generalized cones" (Marr, 1982, p. 303), on the basis of which object recognition is achieved by means of a system of indexing which allows a "newly derived description to be associated with a description [already stored] in the collection." (p. 318) Appropriate perceptually guided behaviours could then be initiated.

From this perspective, the increasing interest in connectionist networks can be understood as arising from the fact that such networks have proved to be superior (in a quantitative sense, as noted below) to rule-based symbol-systems in terms of achieving the dimensional reduction necessary for pattern recognition. Moreover, within mainstream cognitive science, the challenge of explaining object recognition and perceptually guided behaviour is widely assumed (perhaps rightly) to be essentially the same in principle as that faced by the wider AI community, say, in designing pattern recognition systems capable of putting captions to images. And with respect to the latter, the recent advances using multi-layer "deep" networks (Neilson, 2015) (Heaton, 2015) have been impressive. See for example, the "ImageNet computer vision challenge." (Linn, 2015) However, although these developments might lend support to the claim that objection recognition and perceptually guided behaviour can be explained in terms of representations which are second-order isomorphic to their targets, they have no bearing on role A.

The distinction between the two roles is further demonstrated in terms of Marr's three levels of description. The role B requirement on a cognitive system which is to be capable of initiating appropriate perceptually guided behaviours can be specified purely at the top level at which "the system is characterised as a mapping from one kind of information to another" (Marr, 1982, p. 24), that is computing a particular function.[179] The descriptions at the lower levels of

---

[179] Nielsen provides examples of such mappings in discussing the capability of neural networks. "Almost any process you can imagine can be thought of as function computation. Consider the problem of naming a piece of music based on a short sample of the piece. That can be thought of as computing a function. Or consider the

representation and physical realisation are significant only in as much as they impact on the system's ability to meet the top level requirement, that is to provide a particular mapping between sensory inputs and appropriate behaviours. However the requirements of role A impact directly on Marr's middle level description. Now there is a requirement with respect to the mapping between sensory inputs and representations, rather than sensory inputs and perceptually guided behaviours.

Although Edelman is not concerned with positing representations to meet the requirements of role A, he clearly recognizes that second-order isomorphic representations might be problematic with respect to accounting for the richness of perceptual experience:

> … a shift towards the view of representation *of* similarity carries with it a price. The standard version of the problem of qualia actually seems to be exacerbated: on the face of it, it is more difficult to explain the apparent richness of the perceived world if one denies that the shape of each of the constituent objects is fully represented. (1998, p. 464)

Edelman suggests that a "partial solution" might be found if we accept, as suggested in recent discussions of change blindness, that the seeming "richness of the perceived world is, to a considerable extent, apparent." However he notes that difficulties remain[180] and includes an appendix outlining a further "not impossible" solution to the problem. (1998, p. 464) Edelman clearly recognizes that there is a significant issue, with respect to "richness", which remains to be resolved. Moreover, as we will see in Section 4.6, he would give short shrift to the claim that perception is in any significant sense direct.

---

problem of translating a Chinese text into English. Again, that can be thought of as computing a function." (Neilson, 2015)

[180] "… we are still left with the need to explain why and how a tomato looks round and red to the observer, who represents directly only the differences between tomatoes and, say, pears and oranges (as opposed to the shape and the color of the tomato)." (1998, p. 464)

Although the representations discussed in Edelman and Churchland are similar, they differ from Marr's 2½-D sketch. However although the former may be superior in terms of meeting the requirements of role B, there is no reason to suppose that they are any better suited to role A. Moreover, in view of the tension between the two sets of requirements, we might expect the reverse. Indeed, in terms of role A, although it fails to make the grade, the 2½-D sketch appears to be the best of a bad bunch.[181] Unsurprising perhaps, since, as we saw in Section 4.1, it has been put forward as the likely candidate.

At this point we could conclude that the requirements on representations in the two roles are not only different but mutually exclusive, and thus consolidate our argument that, although well suited to role B, the representations posited, or in prospect, within cognitive science cannot meet the requirements of role A. However it might be objected that we have only considered three particular cases and our generalization is unwarranted; indeed we outlined a wider range of representations in Section 3.2.3. Nevertheless, although there may of course be entirely unforeseen developments, there are good reasons to suppose that our conclusions can be extended to any cognitive system which operates in accordance with our current understanding of computation and physical science.

As we noted in Section 3.2.2, although not entirely uncontroversial, the broad consensus within computer science is that, in terms of the functions that can be computed given sufficient time and physical resources, any system that can compute is a Turing machine equivalent. Furthermore, in accordance with this overarching universality claim, it has been proved by Cybenko (1989) and others that single-hidden-layer feed-forward neural networks are "universal approximators" and

---

[181] As a representation of the layout of surfaces in the physical world, the 2½-D sketch is a prime example of an information-rich objective representation, that is a stored description of the environment which is independent of any cognitive agent's sensory apparatus, goals and repertoire of behaviours. Indeed Marr describes it as "an internal representation of objective physical reality." (1982, p. 269)

"universal classifiers". (Hassoun, 1995, p. 46) More loosely put, such networks, of which the Cottrell-type network is one, can in principle "produce (or approximate) any output from any input as long as it has enough hidden neurons in a single layer." (Heaton, 2015, p. 8) For our purposes, the key point is that the computational capabilities, in terms of mapping inputs to outputs, of our example of the Cottrell network cannot be exceeded by the more complex, and perhaps more brain-like, "deep" multilayer networks which have now been developed; though the latter may greatly improve[182] on single-hidden-layer networks in terms of trainability and neuron count. Thus we need not be embarrassed by continuing to draw on such a simple and somewhat dated example in Section 4.4.[183]

Nevertheless, universality with respect to the functions that can be computed by different computational systems does not guarantee universality with respect to the type of representations that can be instantiated therein. For example, it is plausible to suppose that different types of representations might differ in their ability to meet the "certain further conditions" under which, according to Strong Representationalism, "phenomenal character is one and the same as representational content." (Tye, 2009, p. 256) Here an important issue appears to be the accessibility of the content, which might be taken to depend on how the representation is instantiated at the hardware level: analogue v digital, distributed v local etc. On the other hand, it is much less plausible to suppose that representations which differ at the hardware level thereby differ in their ability to meet the requirements of role A. Now the issue is content rather than the instantiation thereof. Thus, it appears that, although certain developments in cognitive science might be taken to lend support to representationalist theories of consciousness, they do not provide candidates for role A and provide no ammunition for a defence of direct realism.

---

[182] The "improvements" may be such that a computation which would require an infeasible number of neurons in one network architecture might be feasible in another.

[183] Indeed we are in the company of Churchland in his recent monograph (2012, p. 62).

This distinction is brought out in the work of O'Brien and Opie on computation, representation and phenomenal consciousness. They maintain that the best prospect for a naturalistic explanation of phenomenal consciousness is to be found by bringing together the "computational theory of mind" and the "representationalist theory of consciousness". (2001) Moreover, they argue that although this explanation faces serious objections in the context of a "classical computational theory of mind"[184], these objections can be resolved within a connectionist theory. As we noted in Section 3.2.3, they maintain that there is a fundamental difference in the nature of representations in analogue and digital computation. And, it is on the basis of this distinction that they defend the representationalist theory of consciousness. Moreover, they could be seen as explicating the further conditions under which, according to Strong Representationalism (Tye, 2009)[185], we become conscious of some but not other representational content, in terms of "activation level representation" and "connection weight representation" respectively.[186]

However, although O'Brien and Opie are defending a representationalist theory of consciousness on the basis of connectionist cognitive science, they are certainly not defending the notion of direct perception or positing representations that meet the requirements of role A. After developing a theory of second-order resemblance, again illustrated by Cottrell's face-recognition network, they conclude that:

> … although our conscious experiences seem to place us in direct contact with the world outside our brains, they actually do no such thing. Conscious experiences are really internal models of the world. … From this perspective, the world of experience—our rich

---

[184] The representationalist theory cannot account for: (1) the distinction between representational content of which we are, and are not, conscious, and (2) the causal efficacy of representations - this requires that representational content must be an intrinsic property of "the brain's representing vehicles".

[185] Tye proposes such further conditions in his PANIC theory: the content must be Poised, Abstract, Non-conceptual, Intentional Content. (1997, p. 137)

[186] Tononi's recent "Integrated information theory of consciousness" might also offer a basis for making this distinction. (2012)

and highly structured phenomenology of body and physical environment—is a "virtual reality" constructed by the brain and "projected outwards" (O'Brien & Opie, 2001)

At this point we rest our case. The representations posited, or in prospect, within cognitive science cannot supply the representational content required of a representation in a direct realist theory of perception. Our conclusion would not trouble mainstream representation-friendly cognitive scientists since they reject the notion of direct perception, and the minority who do accept the notion are representation-averse. Moreover representation-averse disjunctivist direct realist philosophers would be untroubled; they warned against talk of representations. Only representation-friendly representationalist direct realists need be concerned. However, given the importance of representationalism in the late twentieth century rehabilitation of direct realism, our claim is highly significant. Moreover we are now able to reject representationalist/intentionalist direct realism as the last hope of avoiding the argument from causation presented in Section 1.3.

4.6 Locke, Berkeley and the Churchlands' "slippery slope"

For the purpose of defending our central thesis, we claim only that the theories of perception posited within mainstream representation-friendly cognitive science are incompatible with the metaphysical directness-claim. However, in this section we go further to voice concerns about the coherence of such theories. Although beyond our remit, it would be a significant omission not to sketch out these concerns.

We maintained above that our conclusion would not trouble proponents of mainstream cognitive science, since they are primarily concerned with explanations of perceptually guided behaviour (rather than perceptual experience) and are, in any case, not in the business of defending direct perception. However, it is not at all clear where they would position themselves with respect to the metaphysical issues which have occupied Western Philosophy since the Early Moderns.

Where explicit statements are made, eg Smythies and Ramachandran (1997), they do appear to reject the metaphysical directness-claim[187]. Moreover, there are position statements, as in Obrien and Opie quoted above and Edelman below, which appear to involve clear but implicit rejections of the metaphysical directness-claim.[188] Nevertheless, many of those working in cognitive science are seemingly unconcerned with, or avoid, the metaphysical issues. Chemero (2009, p. 185) talks of those "who don't want to be bothered by metaphysical concerns" and cites Clark (1997) "along with nearly every nonphilosopher." Significantly, Austin's dismissal of such concerns is still influential in some quarters. For example, Schwartz takes Austin as sanctioning his, one time, avoidance of the metaphysical issues (2004, p. 255)[189], and note Marr's reference to Austin below.

Although Marr's theory, in terms of information processing and a sequence of three representations (primal sketch, 2½-D sketch and 3-D model representation), might be taken as an implicit endorsement of indirect realism, it seems that he would be unhappy with this interpretation. In outlining the range of considerations that must be addressed before we can be said to have an understanding of perception, he says:

---

[187] In a paper proposing an empirical refutation of direct realism, they are clearly targeting this identity claim. Here they are presenting evidence against "the Direct Realist theory of perception which states that, in vision, the visual field contains the physical object itself, and thus the phenomenal object is identical to the physical object." (Smythies & Ramachandran, 1997, p. 437)

[188] This is further exampled by Gregory. As he puts it: "The essential point about full perception is that indirectly it's related to the external world. We seem to be directly aware of things around us, almost in the world itself continuously. Actually the brain is locked away in its little black box, not of course receiving any light for example directly …all that the brain gets are signals like Morse code, dots and dashes if you like, from the senses, which it has to decode. It has to interpret …. and it constructs the world around us, it's really constructed by an active brain." (Gregory, 2005) This is taken from a radio discussion and may be unscripted. However, for this reason it probably reveals Gregory's underlying position well, without the qualifications and reservations of a formal paper. We find very much the same story in recent popular science documentaries on perception and related issues. For example, the BBC's "What is Reality?" (Eagleman, 2016)

[189] As Schwartz puts it in a paper discussing the ongoing influence of Austin: "my interests have always been more in the psychophysics of perception than in its metaphysics. What I liked about Austin was that he seemed to give me permission to think about the former while allowing me, in good conscience, to avoid the latter." (Schwartz, 2004, p. 255)

> First, and I think foremost, there is the perspective of the plain man. He knows what it is like to see, and unless the bones of one's argument and theories roughly correspond to what this person knows to be true at first hand, one will probably be wrong (a point made with force and elegance by Austin, 1962). (Marr, 1982, p. 4)

It is unclear just what Marr takes the plain man to know.[190] However his seeming endorsement of our pre-critical intuitions and his reference to Austin warns us against characterizing Marr as an indirect realist. Unfortunately, Marr has little to say about his metaphysical assumptions and his position remains unclear.[191]

Although O'Brien and Opie make no explicit reference to direct or indirect realism, as we saw above, their talk of "conscious experience" as a "virtual reality" (2001), would appear to endorse the latter. Similarly, Edelman is clearly sympathetic to indirect realism. In the target paper discussed above, he opens the discussion of "Philosophical implications" by drawing parallels between "Locke's conformity and Shepard's second-order isomorphism"[192] and notes that Locke was "among the first to fully realize the infeasibility of Aristotelian representation by resemblance." (1998, p. 462) Moreover, he notes:

---

[190] Maund suggests that Marr misrepresents the "plain man". (Maund, 2003, p. 18)

[191] "From a philosophical point of view, the approach that I describe is an extension of what have sometimes been called representational theories of mind. On the whole, it rejects the more recent excursions into the philosophy of perception, with their arguments about sense data, the molecules of perception, and the validity of what the senses tell us; instead this approach looks back to an older view, according to which the senses are for the most part concerned with telling one what is there. Modern representational theories conceive of the mind as having access to systems of internal representations; mental states are characterized by asserting what the internal representations currently specify, and mental processes by how such internal representations are obtained and how they interact." (Marr, 1982, p. 6) This brief position statement is unclear on three key points: (i) his use of the highly ambiguous term "representational theories of mind" which can be taken to mean indirect realism or the representationalist (aka intentionalist) version of direct realism, (ii) "the philosophy of perception which he rejects" and (iii) the unspecified "older view" to which he returns.

[192] There are indeed such parallels. According to Locke in his *Essay Concerning Human Understanding*, "Our knowledge, therefore is real only so far as there is a conformity between our ideas and the reality of things." (Locke, 1967, p. 499) Similarly, for Edelman perception is veridical only so far as there is a second-order isomorphism between our representations and distal objects. However, Edelman seems to ignore the fact that in Locke there are both primary and secondary qualities and that the former appear to involve a "conformity" which goes beyond second-order isomorphism. Indeed it is with respect to his espousal of primary qualities, that the Churchlands take issue with Locke, and thereby claim to avoid the slippery slope. (2002)

> ... if second-order isomorphism can be made to work, Locke's "conformity" acquires a new concrete meaning: the order and the connection of ideas is identical to the order and the connection of things. (Edelman, 1998, p. 463)

Although Edelman appears to endorse "indirect realism", he does not use the term and his position with respect to the triad of theories (direct realism, indirect realism and phenomenalism) is never explicitly stated.[193] However, his somewhat cryptic comments in the introduction to his monograph *Representation and Recognition in Vision* (1999) are revealing:

> My long-range goal in raising the issue of veridicality and attempting to treat it formally [by positing and analysing his theory] is to help reinstate it as a *comme il faut* concept – a status which it appears to **have lost between Locke's *Essay* and Berkeley's *Treatise*.** [emphasis added] For ages, veridicality has been a charged term, one whose mention may make some of my philosophically minded readers try to ambush me at every turn of the road. (Edelman, 1999, p. xvi)

Edelman might be understood as recommending Locke's indirect realism, whilst seeking to resist Berkeley's objections thereto. One point is clear. As is evidenced in his more recent monograph *Computing The Mind,* he would reject outright the claim that perception is in any significant sense direct:

> Clouds of points in multidimensional representation spaces are the only entities connected to the external world that a cognitive system can ever get to know. Such a space … presents an unavoidable bottleneck through which any perceptual information must pass on its way to the rest of the brain. (Edelman, 2008, p. 92)

Here Edelman could be seen as setting himself up for, as he puts it, an "ambush" by his "philosophically minded readers". Indeed his theory is clearly exposed to Armstrong's objection to indirect realism, "that if the theory were true, we could have no good reasons for postulating the existence of physical objects to be the causes of our immediate perceptions" (1961, p. 142). This is the Berkeleian challenge to Lockean indirect realism (p. 30).

---

[193] The terms "direct realism", "indirect realism" and their cognates do not appear in the target paper (Edelman, 1998)or the two monographs (Edelman, 1999) (Edelman, 2008).

Again we are reminded of Strawson's warning that, even when we are "trying to entertain a Lockian or scientific realism our language and our natural ways of thinking keep pulling us back to a more primitive view". (1979, p. 53) It seems that whilst Edelman (2008, p. 92) talks in terms of a cognitive system that can only ever get to know clouds of points in multidimensional representation spaces, he keeps being pulled back to a more primitive view, from which he sees himself (qua a cognitive system) in the position of the designer in the engineering example who has direct epistemic access to the targets of the representational structures that she is designing. Edelman appears to have overlooked the "unavoidable bottleneck through which any perceptual information must pass."

Edelman's admission of an "unavoidable bottleneck" (2008, p. 92) is particularly inviting of Armstrong's objection to indirect realism. However the same objection could also be directed at the theories of perception posited by Marr[194] and Churchland as discussed in Sections 4.3 and 4.4.

The Churchlands would strenuously defend their approach against such objections. However, they clearly recognize the danger of the Berkeleian challenge to the unwary cognitive scientist in their "slippery slope" illustration, which appears both in *Neural worlds and real worlds* (2002, p. 904) and *Plato's Camera* (2012, p. 134). Here, urged on by Galileo, Locke puts his foot on the slippery slope only to find himself sliding, despite the best efforts of Kant, to the bottom to join Berkeley and Hegel. Thus, as they see it, Lockean indirect realism is to be avoided, lest the inexorable descent to idealism.

---

[194] Marr, similarly appears to put himself in the position of the designer in the engineering example, and to suppose that he has direct epistemic access to the targets of the representations that he posits. In terms of the top level of description of an information processing device as "a mapping from one kind of information to another" (1982, p. 24), applied to the perceiving cognitive agent, he takes the external world inputs and outputs to be knowns. (p. 41) The challenge, then, is to provide the middle level description of the representations, say the 2½-D sketches, which correspond to these known inputs. To say the least, this seems to be epistemologically dubious.

However, the Churchlands' warning with respect to Locke's precarious position at the top of the slippery slope is certainly not an endorsement of direct realism. They reject "Naive Realism", of which they say, this view "has to ignore the gathering evidence from cognitive science, and neuroscience concerning the brain's constructive processes." (2002, p. 905) Unfortunately this seemingly clear position statement is compromised by the ambiguity in the term "Naïve Realism" and their failure to distinguish between metaphysical and epistemological claims. Thus it is unclear exactly what position they are rejecting, and where they would stand on the ground between an extreme version of "Naïve Realism" [195] and Lockean indirect realism.

We need to understand more clearly why the Churchlands take Locke's position to be so precarious. The source of the problem is Locke's distinction between "real-world properties (primary qualities) and brain-constructed properties (secondary qualities)" (2002, p. 904), and Berkeley's recognition that the arguments which show secondary qualities to be "nothing but mind created responses to a real world" could equally well be applied to primary qualities. Hence the slide into idealism, as the defence of primary qualities succumbs to the Berkeleian objection.

As the Churchlands see it, having rejected "Naïve Realism", the only other response to the Berkeleian challenge is to refuse to embark on the impossible task of defending the notion of primary qualities, that is trying to "establish the resemblance between individual primary qualities and individual properties in the real world". Rather, we should take a stand on "higher-order, multidimensional 'resemblances.'" (p. 907) As they see it, but in our terminology, it is the fruitless quest for first-order isomorphic representations which precipitates the descent into

---

[195] They describe "Naïve Realism" as "boldly claim[ing] that there is a perfect match between appearance and reality". (Churchland & Churchland, 2002, p. 905) However the broad thrust of the paper suggests that they are targeting a less extreme and more plausible version of direct realism.

idealism, whereas second-order isomorphic representations are there to be found instantiated in the brains of cognitive agents and serve as "models for real world survival". We are enjoined to be satisfied and to look no further[196], lest we find ourselves at the bottom of the slippery slope.

However this appears to be a way of avoiding the problem rather than of resolving it. Indeed, they accept that "some contemporary neuroscientists consider the primary-secondary distinction to be unavoidable and the slippery slope to Idealism to be inevitable." (2002, p. 904) Moreover, we would add, the distinction, though not the idealist implication, is widely held in physical science. Thus a major concern remains. That is whether the Churchlands can, whilst being consistent with their theory of perception in terms of representations in vector/parameter space, justify their refusal to join Locke and the indirect realists at the top of the slope.

We suggest that the Churchlands are hoist by their own petard. They, along with rest of mainstream representation-friendly cognitive science (Marr and Edelman included), stand precariously at the top of their slippery slope and cannot in good conscience refuse to take the first step.

4.7 Conclusions

1. Representationalist direct realism must address two issues in a manner which is consistent with the metaphysical directness-claim: the relationship between the world and the representation thereof, and the relationship between the representation and the phenomenal character of perceptual experience. We grant, for the sake of argument, that the second issue can be resolved and focus on the first which has received insufficient attention in the philosophical literature defending direct realism. Representations in the abstract are widely

---

[196] As they put it in the final sentence of their paper: "The reality-appearance distinction ultimately rest on comparisons between the predictive merits of distinct representational models, and the best explanation for why one theory out-predicts another is that one theory is closer to the truth than the other." (2002, p. 907)

discussed, but the question of what representations in cognitive systems actually are, what they represent and how they represent is rarely addressed.

2. Drawing on the representations fleshed-out within cognitive science, we go beyond the standard causal picture to present a more sophisticated causal argument against representationalist direct realism which does not depend on the notion of an immediate object of awareness and thus the Weak Phenomenal Principle. To this end, we distinguish between the requirements on representations in two markedly different roles: role A, to supply the representational content required of a representation in a direct realist theory of perception, and role B, to provide the basis for an explanation of perceptually guided behaviour.

3. Although in principle well suited to role B, none of the representations posited, or in prospect, within cognitive science can meet the requirements of role A. In all cases the representations are second-order isomorphic to their targets, that is representation *by* similarity rather than *of* similarity, and, as such, cannot supply representational content which represents objects as having the properties that they (the objects) actually have. (We make no positive claims as to the type of representation, if any, which would meet the requirements of role A.)

4. None of the theories of perception so far posited, or in prospect, within mainstream representation-friendly cognitive science is compatible with the metaphysical directness-claim. Thus, although representationalist direct realism may be plausible in the context of a loosely formulated notion of representation in the abstract, it cannot be reconciled with the second-order isomorphic relationship between the world and the representation thereof which

appears to be an inherent feature in any representation-friendly theory of perception that could be posited within the naturalistic framework of contemporary natural science.

5. We are now able to reject representationalist/intentionalist direct realism as the direct realist's last hope of avoiding the argument from causation presented in Section 1.3.

5 Summary of conclusions

This dissertation draws on the twenty first century alliance between philosophy and cognitive science to defend the thesis that: *Metaphysical direct realism is incompatible with an explanation of the perceptual process in terms of a naturalistically understood causal chain of events.* In support of this thesis we have drawn the following conclusions:

1. Although causal arguments against direct realism have been dismissed lightly as irrelevant or confused in much of the philosophical literature, such objections arise from a failure to distinguish between metaphysical and epistemological claims, and between *factual* and *strictly-conceptual claims.* Thus sound arguments are rejected along with the patently unsound.

2. When the argument from causation is formalized as an argument against the conjunction of the metaphysical directness-claim *factually* construed and the standard causal picture, it presents a particularly serious objection to the direct realism of contemporary analytic philosophy. Since only the veridical case is at issue, the disjunctivist's objection is avoided altogether and the representationalist's objection is substantially weakened. However, the argument can be challenged by both the small but influential minority of representationalist direct realists, who are prepared to reject the Weak Phenomenal Principle outright, and others who, whilst accepting the Principle, would question the argument's call on the notion of an immediate object of awareness. To break the impasse, we turn to cognitive science to understand more clearly what representations actually are, what they represent and how they represent.

3. Although it has been maintained that "direct realism" is a *strictly-conceptual* thesis which stands apart from the *factual* theses properly addressed within cognitive science, such a deflationary construal would not do justice to the substantial and significant thesis which has been debated since the Early Moderns and is now defended within contemporary analytic philosophy. Moreover such a demarcation would be impossible to maintain in the context of work on embodiment, computation and representation which is undertaken within a twenty first century alliance of the science and philosophy of perception.

4. Mainstream representation-friendly cognitive science fleshes-out the standard causal picture, and thereby provides the basis for a more sophisticated causal argument against representationalist direct realism, which does not depend on the Weak Phenomenal Principle or the notion of an immediate object of awareness. None of the representations so far posited, or in prospect within a naturalistic framework, can supply the representational content required of a representation in representationalist direct realism. Such representations (those posited or in prospect) are second-order isomorphic to their targets and, as such, are well suited to the explanation of perceptually guided behaviour in terms of surrogate reasoning. However they cannot supply representational content which represents objects as having the properties that they (the objects) actually have. Thus mainstream representation-friendly cognitive science is shown to be incompatible with the metaphysical directness-claim.

5. Gibsonian, representation-averse, cognitive science rejects the standard causal picture, thereby disarming the argument from causation, and cannot be shown to be incompatible with the metaphysical directness-claim. Thus, it could be maintained that by virtue of positing a non-standard, but nevertheless naturalistic, causal picture, it does present an

objection to our central thesis. We would reject this suggestion on two counts. Firstly there are, as yet, no fleshed-out representation-averse explanations, and arguably none in prospect. Secondly, although Gibsonian cognitive science might be compatible with the metaphysical directness-claim, it is, arguably, not robustly realist; indeed some of its proponents are avowedly antirealist. Although, outside the scope of this dissertation, we would go further to maintain that representation-averse cognitive science is, in principle, incompatible with the realism-claim. We suggest, as the basis for further work, such an argument in Appendix A.

6. In terms of our central thesis, we have shown that metaphysical direct realism's directness-claim is incompatible with an explanation of the perceptual process in terms of mainstream representation-friendly cognitive science. Moreover, if the argument proposed in Appendix A goes through, neither mainstream nor Gibsonian representation-averse cognitive science, is compatible with a theory of perception which is at the same time both direct and realist in the sense that metaphysical direct realism requires.

References

Armstrong, D. M., 1961. *Perception and the Physical World.* London: Routledge & Kegan Paul.

Arnau, E., Estany, A., Gonzalez del Solar, R. & Sturm, T., 2014. The extended cognition thesis: Its significance for the philosophy of (cognitive) science. *Philosophical Psychology*, 27(1), pp. 1-18.

Austin, J. L., 1962. *Sense and Sensibilia.* 1 ed. London Oxford New York: Clarendon Press.

Aydede, M., 2013. Pain. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Spring 2013 Edition).* s.l.:s.n.

Bechtel, W., Mandik, P., Mundale, J. & Stufflebeam, R. S., 2001. *Philosophy and the Neurosciences.* s.l.:Blackwell.

Bennett, M. & Hacker, P., 2011. Reply to a critical review by Professor John Smythies. *Perception,* 40(2), pp. 249 - 251.

Bickhard, M. & Richie, D. M., 1983. *On the Nature of Representation: A Case Study of James Gibson's Theory of Perception.* s.l.:Praeger.

Bonjour, L., 2013. Epistemological Problems of Perception. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Spring 2013 Edition).* s.l.:s.n.

Brewer, B., 2004. Realism and the Nature of Perceptual Experience. *Philosophical Issues - A Supplement to NOUS,* 14(1), pp. 61 - 77.

Brewer, B., 2007. Perception and its objects. *Philosophical Studies,* Volume 132, pp. 87-97.

Brooks, R., 1991. Intelligence without representation. *Artificial Intelligence,* 47(1-3), pp. 139 - 159.

Bruce, V., Green, P. & Georgeson, M., 2003. *Visual Perception.* s.l.:Psychology Press.

Byrne, A., 2001. Intentionalism Defended. *Philosophical Review,* Volume 110, pp. 199 - 240.

Byrne, A. & Logue, H., 2008. Either/Or. In: A. Haddock & F. Macpherson, eds. *Disjunctivism: Perception, Action, Knowledge.* s.l.:Oxford University Press.

Chemero, A., 1998. A Stroll Through The Worlds of Animats and Humans: Review of Being There: Putting Brain, Body and World Together Again by Andy Clark. *Psyche,* Volume 4.

Chemero, A., 2009. *Radical Embodied Cognitive Science.* s.l.:Bradford Books.

Chemero, A. & Silberstein, M., 2008. *Defending Extended Cognition.* s.l., s.n., pp. 129-134.

Churchland, P. M. & Churchland, P. S., 2007. *Neurofilosophia,* s.l.: s.n.

Churchland, P. M., 2012. *Plato's Camera - How the Physical Brain Captures a Landscape of Abstract Universals.* s.l.:MIT Press.

Churchland, P. M. & Churchland, P. S., 1990. Could a Machine Think?. *The Scientific American*, January, pp. 26 - 31.

Churchland, P. S., 2002. *Brain-Wise Studies in Neurophilosophy.* s.l.:MIT Press.

Churchland, P. S., 2004. *Brain-Wise: An interview with Patricia Smith Churchland* [Interview] (1 March 2004).

Churchland, P. S. & Churchland, P. M., 2002. Neural worlds and real worlds. *Nature Reviews - Neuroscience,* pp. 903 - 907.

Clark, A., 1997. *Being There - Putting the Brain, Body, and World Together Again.* s.l.:Bradford Books.

Clark, A. & Chalmers, D., 1998. The Extended Mind. *Analysis,* 58(1), pp. 7-19.

Copeland, J., 1993. *Artificial Intelligence: A Philosophical Introduction.* s.l.:Blackwell.

Copeland, J., 2004. Computation. In: L. Floridi, ed. *The Blackwell Guide to the Philosophy of Computing and Information.* s.l.:Blackwell, pp. 3 - 18.

Copeland, J. & Shagrir, O., 2007. Physical Computation: How General are Gandy's Principles for Mechanisms?. *Minds and Machines,* 17(2), pp. 217 - 231.

Crane, T., 2005. What is the Problem of Perception. *Synthesis Philosophica.*

Crane, T., 2011. The Problem of Perception. In: E. Zalta, ed. *The Stanford Encyclopedia of Perception (Spring 2011 Edition).* s.l.:s.n.

Crane, T. & French, C., 2015. The Problem of Perception. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Winter 2015 Edition).* s.l.:s.n.

Crick, F., 1994. *The Astonishing Hypothesis.* s.l.:Simon and Schuster.

Crooks, M., 2008. The Churchlands' War on Qualia. In: E. L. Wright, ed. *The Case for Qualia.* s.l.:MIT Press.

Cucen, K., 1998. *Heidegger's Reading of Descartes' Dualism: The Relation of Subject and Object.* Boston, s.n.

Cummins, R., 1991. The Role of Representations in Connectionist Explanations of Cognitive Capacities. In: W. Ramsey, S. Stich & D. Rumelhart, eds. *Philosophy and Connectionist Theory.* s.l.:Lawrence Erlbaum Associates, Inc, pp. 91 - 114.

Cummins, R., 1996. *Representations, Targets and Attitudes.* s.l.:MIT Press.

Cybenko, G., 1989. Approximations by superposition of sigmoidal functions. *Mathematics of Control, Signals and Systems,* 2(4), pp. 303-314.

Dancy, J., 1985. *Introduction to Contemporary Epistemology.* s.l.:Blackwell.

Dawkins, R., 2006. *The God Delusion.* s.l.:Bantam Press.

Dennett, D., 1993. *Consciousness Explained.* s.l.:Penguin Books.

d'Espagnat, B., 1979. The Quantum Theory and Reality. *Scientific American*, November.

Dictionary, A. H., 2016. *The American Heritage Dictionary of the English Language.* 5 ed. s.l.:Houghton Mifflin Company.

Dretske, F., 1988. *Explaining Behaviour.* s.l.:MIT Press.

Dretske, F., 1990. Seeing, Believing, and Knowing. In: D. Osherson, S. Kosslyn & J. Hollerbach, eds. *Visual Cognition and Action, vol II.* s.l.:MIT Press, pp. 129-48.

Dretske, F., 1995. The Oxford Companion to Philosophy. In: T. Honderich, ed. s.l.:Oxford University Press, p. 602.

Dretske, F., 1997. *Naturalizing the Mind.* s.l.:MIT Press.

Dretske, F., 1999. *Knowledge and the flow of information.* s.l.:CSLI Publications.

Dretske, F., 2000. *Perception Knowledge and Belief.* s.l.:Cambridge University Press.

Dretske, F., 2003. Experience as Representation. *Philosophical Issues,* Volume 13, pp. 67-82.

Dretske, F., 2004. Seeing, Believing, and Knowing. In: R. Schwartz, ed. *Perception.* s.l.:Blackwell, pp. 268 - 286.

Dreyfus, H., 2005. Merleau-Ponty and Recent Cognitive Science. In: T. Carman & M. Hansen, eds. *The Cambridge Companion to Merleau-Ponty.* s.l.:Cambridge University Press, pp. 129 - 150.

Dreyfus, H., 2005. Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise. *Proceedings of the American Philosophical Association,* 79(2), pp. 47 - 65.

Eagleman, D., 2016. *"What is Reality".* [Online, BBC documentary] [Accessed February 2016].

Edelman, G., 1992. *Bright Air, Brilliant Fire - On the Matter of the Mind.* s.l.:Penguin Books.

Edelman, S., 1998. Representation is representation of similarities. *Behavioural and Brain Sciences,* Volume 21, pp. 449-498.

Edelman, S., 1999. *Representations and Recognition in Vision.* s.l.:MIT Press.

Edelman, S., 2008. *Computing the Mind: How the Mind Really Works.* s.l.:Oxford University Press.

Einstein, A., Podolsky, B. & Rosen, N., 1935. Can Quantum-Mechanical Description of Physical Reality be Considered Complete?. *Physical Review,* Volume 47.

Feynman, R., 1963. *The Feynman Lectures on Physics.* s.l.:Addison-Wesley.

Fish, W., 2004. The Direct/Indirect Distinction in Contemporary Philosophy of Perception. *Essays in Philosophy.*

Fish, W., 2009. *Perception, Hallucination and Illusion.* s.l.:Oxford University Press.

Fish, W., 2010. *Philosophy of Perception: A Contemporary Introduction.* s.l.:Routledge.

Fodor, J. & Pylyshyn, Z., 2002. How Direct Is Visual Perception?. In: A. Noe & E. Thompson, eds. *Vision and Mind.* s.l.:MIT Press, pp. 167 - 227.

Forrest, P., 2012. The Identity of Indiscernibles. In: E. N. Zalta, ed. *The Stanford Encyclopedia of Philosophy.* s.l.:s.n.

Foster, J., 2000. *The Nature of Perception.* s.l.:Oxford University Press.

Fumerton, R., 1985. *Metaphysical and Epistemological Problems of Perception.* s.l.:University of Nebraska Press.

Fumerton, R., 2005. *Direct Realism, Introspection and Cognitive Science.* s.l., s.n.

Fumerton, R., 2006. Philosophy, Introspection and Cognitive Science. *Philosophy and Phenomenological Research,* 73(3), pp. 680 - 695.

Gellner, E., 1959. *Words and Things.* s.l.:Routledge and Kegan Paul.

Gendler, T. & Hawthorne, J., 2006. *Perceptual Experience.* Oxford: Clarendon Press.

Gibson, J., 1950. *The Perception of the Visual World.* Boston: Houghton Mifflin Company.

Gibson, J., 1967. New Reasons for Realism. *Synthese,* 17(2), pp. 162 - 172.

Gibson, J., 1968. *The Senses Considered as Perceptual Systems.* London: George Allen & Unwin.

Gibson, J., 1974. *Presentation at Ohio State University (sound recording).* [Online]
Available at: https://www.youtube.com/watch?v=hwRxUyuQEgc
[Accessed December 2015].

Gibson, J., 1979. *The Ecological Approach to Visual Perception.* s.l.:Houghton Miflin Co.

Gibson, J., 2002. A Theory of Direct Visual Perception. In: A. Noë & E. Thompson, eds. *Vision and Mind.* s.l.:MIT Press.

Goldman, A., 2015. Reliabilist Epistemology. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy.* Spring 2011 ed. s.l.:s.n.

Goldstein, B., 1981. The Ecology of J. J. Gibson's Perception. *Leonardo,* 14(3), pp. 191 - 195.

Gould, S. J., 1999. *Rocks of Ages.* s.l.:Ballatine Books.

Gregory, R., 1987. Turing, Alan Mathison . In: R. Gregory, ed. *The Oxford Companion to the Mind.* s.l.:Oxford University Press, p. 784.

Gregory, R., 2002. Perceptions as Hypotheses. In: A. Noë & E. Thompson, eds. *Vision and Mind.* s.l.:MIT Press, pp. 111 - 133.

Gregory, R., 2005. *Perception and the Senses (BBC In Our Time).* [Online]
Available at: http://www.bbc.co.uk/programmes/p003k9h0
[Accessed December 2015].

Grice, H. P., 1961. The Causal Theory of Perception. *Proceedings of the Aristotelian Society,* Volume 35, pp. 121 - 152.

Griffiths, P., 2008. *The Case Against Direct Realism - MPhil dissertation (University of Manchester).* s.l.:s.n.

Groblacher, S., 2007. An experimental test of non-local realism. *Nature,* Volume 446, p. 871.

Gross, C., 2002. Genealogy of the Grandmother Cell. *The Neuroscientist ,* 8(1), pp. 85 - 90.

Gross, C. G., 1999. The Fire that Comes from the Eye. *The Neuroscientist,* Volume 5, pp. 58 - 64.

Grush, R., 2008. Representation Reconsidered - Reviewed by Rick Grush. *Note Dame Philosophical Reviews.*

Hacking, I., 1983. *Representing and Intervening.* s.l.:Cambridge University Press.

Hamlyn, D. W., 1961. *Sensation and Perception - A History of the Philosophy of Perception.* London: Routledge & Keegan Paul.

Hassoun, M., 1995. *Fundamentals of Artificial Neural Networks,* s.l.: Bradford Books.

Hatfield, G., 2003. Representation and constraints: the inverse problem and the structure of visual space. *Acta Psychologica,* pp. 355-378.

Haugeland, J., 1990. Representational Genera. In: W. Ramsey, S. Stich & D. Rummelhart, eds. *Philosophy and Connectionist Theory.* s.l.:Lawrence Erlbaum Associates, pp. 61 - 90.

Hawking, S., 2010. *The Grand Design.* s.l.:Bantam Books.

Hayward, W. & Tarr, M., 2005. Visual Perception II: High-Level Vision. In: K. Lamberts & R. Goldstone, eds. *The Handbook of Cognition.* s.l.:Sage, pp. 48 - 70.

Heaton, J., 2015. *Artificial Intelligence for Humans, Volume 3.* s.l.:Heaton Research Inc.

Hirst, R. J., 1959. *The Problems of Perception.* London: George Allen and Unwin Ltd.

Hodgson, D., 1996. Nonlocality, local indeterminism and consciousness. *Ratio.*

Huemer, M., 2001. *Skepticism and the Veil of Perception.* s.l.:Rowman and Littlefield Publishers, Inc.

Huemer, M., 2011. Sense-Data. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition),.* s.l.:Stanford.

Hunt, M., 1994. *The Story of Psychology.* s.l.:Anchor Books.

Hutto, D. & Myin, E., 2012. *Radicalizing Enactivism.* s.l.:MIT Press.

Jackson, F., 1977. *Perception- A representative theory.* s.l.:Cambridge University Press.

Jordan, M., 2001. Recurrent Networks. In: R. Wilson & F. Keil, eds. *The MIT Encyclopedia of the Cognitive Sciences.* s.l.:Bradford Books, pp. 709 - 712.

Kamthane, A., 2007. *Introduction to Data Structures in C.* s.l.:Pearson Education India.

Karakostas, V., 2012. Realism and Objectivism in Quantum Mechanics. *Journal for General Philosophy of Science,* 43(1).

Keijzer, F., 2001. *Representation and Behaviour.* s.l.:MIT Press.

Kelly, S., 2005. Seeing Things in Merleau-Ponty. In: T. Carman & M. Hansen, eds. *The Cambridge Companion to Merleau-Ponty.* s.l.:Cambridge University Press, pp. 74 - 110.

Kind, A., 2007. Restrictions on Representationalism. *Philosophical Studies,* Volume 134, pp. 405 - 427.

Kirk, R., 2015. Zombies. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy.* s.l.:s.n.

Kulvicki, J., 2004. Isomorphism in information carrying systems. *Pacific Philosophical Quarterly.*

Lau, J. & Deutsch, M., 2013. Externalism about Mental Content. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy.* s.l.:s.n.

le Morvan, P., 2004. Arguments against direct realism and how to counter them. *American Philosophical Quarterly,* Volume 41, pp. 221 - 234.

Linn, A., 2015. *Microsoft researchers win ImageNet computer vision challenge,* s.l.: Microsoft.

Locke, D., 1967. *Perception and our Knowledge of the External World.* s.l.:George Allen & Unwin Ltd.

Locke, J., 1967. *An Essay Concerning Human Understanding.* s.l.:Penguin Books.

Lockwood, M., 1989. *Mind, Brain & the Quantum.* s.l.:Basil Blackwell.

Logue, H., 2012. Why Naive Realism?. *Proceedings of the Aristotelian Society ,* Volume CXII, Part 2, pp. 211-237.

Loomis, J., 1993. *Understanding of Synthetic Experience Must Begin with the Analysis of Ordinary Perceptual Experience.* San Jose, IEEE, pp. 54 - 57.

Lycan, W., 2015. Representational Theories of Consciousness. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition).* s.l.:s.n.

Macarthur, D., 2003. Mc Dowell, Scepticism, and the 'Veil of Perception'. *Australasian Journal of Perception,* 81(2), pp. 175 - 190.

Marr, D., 1982. *Vision.* s.l.:W H Freeman and Co.

Marr, D., n.d. *Representing and Computing Visual Information.* [Online]
Available at:
http://mitpress.mit.edu/sites/default/files/titles/content/9780262730594_sch_0001.pdf
[Accessed December 2015].

Martin, M., 1998. Setting Things Before the Mind. In: A. O'Hear, ed. *Current Issues in Philosophy of Mind.* s.l.:Cambridge University Press, pp. 157-180.

Martin, M., 2000. Beyond Dispute. In: T. Crane & S. Patterson, eds. *The History of the Mind-Body Problem.* London: Routledge, pp. 195-231.

Martin, M., 2004. The Limits of Self-Awareness. *Philosophical Studies,* Volume 120, pp. 37 - 89.

Maund, B., 2003. *Perception.* s.l.:Acumen.

Mc Dermid, D., 2004. The Real World Regained? Searle's External Realism Examined. *Kriterion,* Volume 18, pp. 1-9.

McDowell, J., 1982. Criteria, Defeasibility and Knowledge. *Procedings of the British Academy ,* Volume 68, pp. 455-479.

McGinn, C., 1997. Can We Solve the Mind-Body Problem. In: N. Block, O. Flanagan & G. Guzeldere, eds. *The Nature of Consciousness.* s.l.:Bradford Books, pp. 529 - 542.

McLendon, H., 1956. Has Russell Prove Naive Realism Self-Contradictory?. *The Journal of Philosophy,* 53(9), pp. 289-302.

Michaels, C. & Carello, C., 1981. *Direct Perception.* s.l.:Prentice-Hall.

Mundle, C. K., 1959. Common Sense versus Mr Hirst's Theory of Perception. *Proceedings of the Aristotelian Society,* p. 146.

Nakayama, K., 1994. James J. Gibson - An Appreciation. *Psychological Review,* 101(2), pp. 329 - 335.

Nathan, N. M., 2005. Direct Realism: Proximate Causation and the Missing Object. *Acta Analytica,* 20(3), pp. 3 - 6.

Neilson, M., 2015. *Neural Networks and Deep Learning.* s.l.:Determination Press.

Nelson, J., 1964. An Examination of D. M. Armstrong's Theory of Perception. *American Philosophical Quarterly,* 1(2), pp. 154 - 160.

Noë, A., 2002. Direct Perception. In: L. Nagel, ed. *Macmillan Encyclopedia of Cognitive Science.* s.l.:Macmillan.

Noë, A. & Thompson, E., 2002. *Vision and Mind.* s.l.:MIT Press.

O'Brien, G. & Opie, J., 2001. Connectionist vehicles, structural resemblance, and the phenomenal mind. *Communication & Cognition ,* 34(1-2), pp. 13-38.

O'Brien, G. & Opie, J., 2004. Towards a Structuralist Theory of Mental Representation. In: H. Clapin, P. Staines & P. Slezak, eds. *Representation in Mind: New Approaches to Mental Representation.* s.l.:Elsevier.

O'Brien, G. & Opie, J., 2009. The role of representation in computation. *Cognitive Processing,* 10(1).

O'Brien, G. & Opie, J., 2011. Representation in Analog Computation. In: A. Newen, A. Bartels & E. Jung, eds. *Knowledge and Representation.* s.l.:University of Chicago Press.

Oytam, Y. & Neilson, P., 2007. Berkeleian Principles in Ecological Realism: An Ontological Analysis. *Ecological Psychology,* 19(3), pp. 265 - 304.

Palmer, S., 1978. Fundamental Aspects of Cognitive Representation. In: E. Rosch & B. Lloyd, eds. *Cognition and Categorization.* s.l.:Lawrence Erlbaum Associates, pp. 259 - 303.

Pasch, A., 1957. Science, Perception, and Some Dubious Epistemological Motives. *Philosophical Studies,* 8(4), pp. 55-61.

Patton, L., 2014. Hermann von Helmholtz. In: E. Zalta, ed. *Stanford Encyclopedia of Philosophy.* s.l.:Stanford.

Penrose, R., 1989. *The Emperor's New Mind.* s.l.:Oxford University Press.

Penrose, R., 1994. *Shadows of the Mind.* s.l.:Oxford University Press.

Piccinini, G., 2015. Computation in the Physical Sciences. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition).* s.l.:s.n.

Port, R., 2002. The Dynamical Systems Hypothesis in Cognitive Science. In: L. Nadel, ed. *Encyclopedia of Cognitive Science.* s.l.:Macmillan Company, pp. 1027-1032.

Price, H. H., 1950. *Perception.* 2 ed. s.l.:Methuen.

Quian Quiroga, R., Reddy, L., Kreiman, G. & Fried, I., 2005. Invariant visual representation by single neurons in the human brain. *Nature,* Volume 4365, pp. 1102 - 1107.

Quine, W. V. O., 1951. Two Dogmas of Empiricism. *The Philosophical Review,* Volume 60, pp. 20-43.

Ramsey, W., 2001. Connectionism, Philosophical Issues. In: R. Wilson & F. Keil, eds. *The MIT Encyclopedia of the Cognitive Sciences.* s.l.:Bradford Books, pp. 186 - 187.

Ramsey, W., 2007. *Representation Reconsidered.* s.l.:Cambridge University Press.

Reid, T., 1785. *Essays on the Intellectual Powers of Man.* s.l.:Google Books.

Robinson, H., 1994. *Perception.* 1 ed. s.l.:Routledge.

Rock, I., 2004. Inference in Perception. In: R. Schwartz, ed. *Perception.* s.l.:Blackwell, pp. 200 - 216.

Rowland, T. & Weisstein, E., 2015. *"Principle of Computational Equivalence." From MathWorld--A Wolfram Web Resource.,* s.l.: Wolfram Research.

Rupert, R., 2008. Causal Theories of Mental Content. *Philosophy Compass,* 3(2), pp. 353-380.

Russell, B., 1927. *The Analysis of Matter.* s.l.:Kegan Paul.

Ryle, G., 1953. Ordinary Language. *Philosophical Review,* 62(2), pp. 167 -186.

Sanders, J., 1993. Merleau-Ponty, Gibson and the materiality of meaning. *Man and World,* Volume 26, pp. 287-302.

Sanders, J., 1997. An Ontology of Affordances. *Ecological Psychology,* 91(1), pp. 97-112.

Schliesser, E., 2008. Hume's Newtonianism and Anti-Newtonianism. In: E. Zalta, ed. *Stanford Encyclopedia of Philosophy.* s.l.:s.n.

Schwartz, R., 2004. To Austin or not to Austin, that's the disjunction. *Philosophical Studies,* Volume 120, pp. 255 - 263.

Searle, J., 1984. *Minds, Brains and Science.* s.l.:British Broadcasting Organisation.

Searle, J., 1998. *Mind Language and Society.* s.l.:Basic Books.

Searle, J., 1998. *The Mystery of Consciousness.* s.l.:Granta Books.

Searle, J., 2003. *John Searle on Gibson and Direct Perception.* [Online]
Available at: https://www.youtube.com/watch?v=ve0c0B47xJw
[Accessed December 2015].

Searle, J., 2004. *Biological Naturalism,* s.l.: s.n.

Searle, J., 2015. *Seeing Things as They Are.* s.l.:Oxford University Press.

Shapiro, L., 2011. *Embodied Cognition.* London and New York: Routledge.

Sheldrake, R., 2005. The Sense of Being Stared At: Part 2. *Journal of Consciousness Studies,* 12(6), pp. 32 - 49.

Shepard, R. N., 1975. Form, formation and transformation of internal representations. In: R. l. Solso, ed. *Information processing and cognition: The Loyola Symposium.* s.l.:Lawrence Erlbaum Associates.

Shepard, R. N. & Chipman, S., 1970. Second-order isomorphism of internal representations: Shapes of states.. *Cognitive Psychology,* 1(1), pp. 1 - 17.

Sieg, W., 2001. Church-Turing Thesis. In: R. Wilson & F. Keil, eds. *The MIT Encyclopedia of the Cognitive Sciences.* s.l.:Bradford Books, pp. 116 - 117.

Smith, A., 2002. *The Problem of Perception.* s.l.:Harvard University Press.

Smith, A. D., 2002. *The Problem of Perception (Publisher's summary),* s.l.: s.n.

Smythies, J., 2009. Philosophy, perception, and neuroscience. *Perception,* 38(5), pp. 638 - 651.

Smythies, J., 2011. Reply to Bennett and Hacker. *Perception,* Volume 40, pp. 252 - 254.

Smythies, J. R., 1993. The impact of contemporary neuroscience and introspection psychology on the philosophy of perception. In: E. Wright, ed. *New Representationalisms.* s.l.:Avebury, pp. 205 - 231.

Smythies, J. R. & Ramachandran, V. S., 1997. An empirical refutation of the direct realist theory of perception. *Inquiry: An Interdisciplinary Journal of Philosophy,* Volume 40, pp. 437 - 438.

Snowdon, P., 1992. How to interpret "direct perception". In: T. Crane, ed. *The Contents of Experience.* s.l.:Cambridge University Press.

Snowdon, P., 2002. Perception, Vision and Causation. In: A. Noë & E. Thompson, eds. *Vision and Mind.* s.l.:MIT Press, pp. 151 - 166.

Stephanich, L., 1991. Heidegger: Between Idealism and Realism. *The Harvard Review of Philosophy,* Spring.pp. 20 -298.

Stout, G., 1922. Prof. Alexander's Theory of Sense Perception. *Mind,* Volume xxxx, pp. 385 - 412.

Strawson, P. F., 1979. Perception and its objects. In: G. F. Macdonald, ed. *Perception and Identity.* s.l.:Macmillan Press Ltd, pp. 41 - 60.

Stroud, B., 2008. The Charm of Naturalism. In: M. de Caro & D. Macarthur, eds. *Naturalism in Question.* s.l.:Harvard University Press, pp. 21 - 35.

Suchting, W., 1969. Perception and the Time-gap Argument. *Philosophical Quarterly,* 19(74), pp. 46-56.

Swoyer, C., 1991. Structural Representation and Surrogate Reasoning. *Synthese,* 87(3), pp. 449 - 508.

Thompson, B., 2008. Representationalism and the Argument from Hallucination. *Pacific Philosophical Quarterly,* Volume 89, pp. 384 - 412.

Tononi, G., 2012. Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie,* Volume 150, pp. 290 - 326.

Turvey, M., 2003. *Direct perception: A matter of non-locality, self-motion, and impredicativity.* [Online]
Available at: http://caribou.cc.trincoll.edu/depts_ecopsyc/isep/ICPA-conf/
[Accessed December 2015].

Tye, M., 1997. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind.* s.l.:Bradford Books.

Tye, M., 2000. *Consciousness, Color, and Content (Representation and Mind).* s.l.:MIT Press.

Tye, M., 2009. Representational Theories of Consciousness. In: B. McLaughlin, A. Beckermann & S. Walter, eds. *The Oxford Handbook of the Philosophy of Mind.* s.l.:Oxford University Press, pp. 253 - 267.

Tye, M., 2015. Qualia. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Fall 2015 Edition).* s.l.:s.n.

Ullman, 1980. Against Direct Perception. *Behavioral and Brain Sciences,* 3(03), pp. 373 - 381.

Valberg, J. J., 1992. *The Puzzle of Experience.* Oxford: Clarendon Press.

van Gelder, T., 1998. The Dynamical Hypothesis in Cognitive Science. *Behavioral and Brain Sciences,* 21(5), pp. 615 - 665.

van Gelder, T., 2001. Distributed vs Local Representation. In: R. Wilson & F. Keil, eds. *The MIT Encyclopedia of the Cognitive Sciences.* s.l.:Bradford Books, pp. 236 - 238.

van Gelder, T., 2001. Dynamic Approaches to Cognition. In: R. Wilson & F. Keil, eds. *The MIT Encyclopedia of the Cognitve Sciences.* s.l.:Bradford Books, pp. 244 - 246.

Varela, F., Thompson, E. & Rosch, E., 1991. *The Embodied Mind: Cognitive Science and Human Experience.* s.l.:MIT Press.

Vision, G., 1997. *Problems of Vision.* s.l.:Oxford University Press.

von Helmholtz, H., 2004. Treatise on Physiological Optics. In: R. Schwartz, ed. *Perception.* s.l.:Blackwell, pp. 42 - 49.

Wagemans, J., 1986. Direct Theory of Perception an Evaluation by Representatives of Indirect Theories of Perception. *L'Annee Psychologique,* pp. 261 - 273.

Ward, A., 1976. Direct and Indirect Realism. *American Philosophical Quarterly,* 13(4), pp. 287 - 294.

Warnock, G. J., 1967. *The Philosophy of Perception.* s.l.:Oxford University Press.

Weisstein, E. W., 2013. *Isomorphism.* [Online]
Available at: http://mathworld.wolfram.com/Isomorphism.html
[Accessed 9 Oct 2013].

Wheeler, M., 2005. *Reconstructing the Cognitive World.* s.l.:MIT Press.

Williamson, T., 1995. Realism and anti-realism. In: T. Honderich, ed. *The Oxford Companion to Philosophy.* s.l.:Oxford University Press, pp. 746 - 748.

Wilson, R. & Foglia, L., 2011. Embodied Cognition. In: E. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition).* s.l.:s.n.

Winner, G. A. et al., 2002. Fundamentally misunderstanding visual perception: Adult's belief in visual emissions. *American Psychologist ,* 5(6-7), pp. 417 - 424.

Wittgenstein, L., 1953. *Philosophical Investigations.* 2 ed. Oxford: Basil Blackwell.

Wolfram, S., 2002. *A New Kind of Science.* s.l.:Wolfram Media Inc.

Zahidi, K., 2013. Non-representationalist cognitive science and realism. *Phenomenology of the Cognitive Sciences.*

Appendix A: Representation-averse cognitive science and the realism-claim

In Chapter 4 we concluded that mainstream representation-friendly cognitive science is incompatible with direct realism's metaphysical directness-claim. Moreover in Section 3.3 we suggested that representation-averse Gibsonian cognitive science is also incompatible with metaphysical direct realism, but now it is the realism-claim that is at issue. Here we propose, as the basis of further work, an argument to this effect.

 In the main body of the dissertation we defended our central thesis on the basis of the argument from causation. However, in order to address concerns with respect to the realism-claim, we would need to pursue a different line of argument which draws on the ontological implications of representation-averse theories of perception. Section A1 introduces such an argument and emphasizes the difference from that which has gone before. In Section A2 we maintain that the efficacy of perceptually guided behaviour cannot be explained in non-representational terms without positing a meaning laden world or a dissolution of the perceiver/perceived dichotomy. In Sections A3 and A4 we argue that neither of these posits is compatible with the realism-claim. Section A5 summarizes our provisional conclusions along with caveats thereon.

A1 The structure of the argument

The argument presented in Chapter 4 is a sophisticated version of the argument from causation which draws on the standard causal picture, as it has been fleshed-out in the theories put forward by mainstream representation-friendly cognitive science, to reject the metaphysical directness-claim. However, a proponent of representation-averse Gibsonian cognitive science could maintain that there are naturalistic theories of perception which are compatible with the metaphysical directness-claim. In response, we argue that such representation-averse theories,

although they avoid the argument from causation by virtue of rejecting the standard causal picture, are incompatible with the realism-claim.

However, we need to be reminded at the outset that the claim that sophisticated perceptually guided behaviour can be explained without recourse to representations is highly controversial and is rejected within mainstream, non radical, cognitive science. Moreover we accept that Gibsonian representation-averse cognitive science may be compatible with the metaphysical directness-claim, not because it (the cognitive science) provides a fleshed-out theory, but rather, as discussed in Section 3.3, because it defuses the objections to direct perception that are voiced by mainstream cognitive science. Thus we are being somewhat charitable in granting that such representation-averse explanations are possible, and that twenty first century radical embodied cognitive science thereby constitutes a putative counter example to our central thesis which needs to be addressed.

As we saw in Section 3.3.3, representation-averse cognitive science brings the realism issue to the fore. Gibson's seminal work has been linked to the Phenomenological Tradition particularly Merleau-Ponty (Sanders, 1993) and Heidegger (Wheeler, 2005, p. 285), Varela et al reject "realism or subject/objectivism" (1991, p. 9), Oytam and Neilson draw parallels with "the principles of [Berkeleian] idealism" (2007, p. 302), and Chemero talks in terms of "realism about the world as it is perceived and experienced" (2009, p. 150) and makes the connection with the radical empiricism and neutral monism of William James (p. 183). However, although the tension between Gibsonian representation-averse cognitive science and "realism" is widely recognized, our argument is unavoidably less direct than that presented in Chapter 4. There are two issues which impact on the type of argument that can be framed.

Firstly, whereas representation-friendly cognitive science presents relatively detailed explanations of the perceptual process for examination, representation-averse cognitive science

does not. Indeed the holistic and impenetrable nature of non-representational theories of perception appears to be one of their essential features. Compare Gibson's undeveloped notion of information pick up by psychological resonance and his radio analogy (1968, p. 271) with Marr's computational theory with its primal sketch, 2½-D sketch and 3-D model representation (1982). As Shapiro puts it, "as analogies go, resonating has hardly the precision or fecundity of the computational analogies that Gibson wishes to displace." (2011, p. 36)   And, as discussed below, in the context of radical embodied cognitive science, Chemero (2009, p. 31) notes the inherently "nondecompossable" nature of non-linear dynamical systems, whose behaviour cannot be explained reductively in terms of the functionality of their parts.

Secondly, although considerations with respect to the constitution and operation of our, or some other species', perceptual systems might provide the basis for objections to the two directness claims, such considerations have no purchase against the realism-claim. The latter is a claim about how the world is, irrespective of whether, and how so, we or other species perceive it. As discussed in Section 1.3.1, the argument from causation is not an argument against, but rather an argument which assumes, the realism-claim. Arguments which address the realism-claim must take a different form.

With respect to representation-averse theories, we argue that any attempt to account for sophisticated perceptually guided behaviour without recourse to the notion of representation creates an explanatory gap which can only be filled at the ontological expense of positing a meaning laden world or a dissolution of the perceiver/perceived dichotomy, neither of which is compatible with the realism-claim. The further requirement for an explanation of perceptual experience need not be addressed. If perceptually guided behaviour cannot be explained without espousing an anti-realist ontology, our case is made.

Our argument is in two stages: (i) that representation-averse theories cannot account for the efficacy of perception without positing a new ontology, and (ii) that the new ontology is incompatible with the realism-claim. Amongst proponents of Gibsonian cognitive science, the connection between representation and ontology is widely recognized. Writing from the perspective of ecological psychology, Oytam and Neilson emphasize the connection:

> Most important, there appears to be a direct link between underlying ontology and rejection (or acceptance) of the validity of *representation* as a process in the realization of perception-action. Historically those who rejected representation did so on the basis of definite ontological axioms concerning the status of the perceiving-acting being and its environment. (2007, p. 266)

And conclude:

> In ecological psychology, we have found that an anti-representational perspective of perception-action grounded upon an ontology within which the principles of idealism appears consistently to inhere. (p. 302)

However the second step of the argument, from ontology to the rejection of the realism-claim, is more contentious. Much depends on how "realism" is defined. Indeed although Oytam and Neilson are sympathetic to the "principles of idealism", they talk in terms of realism, albeit "ecological realism". Moreover, although Chemero recognizes that radical embodied cognitive science involves "ontological funny business" and "requires a new ontology" (2009, p. 135), he maintains that it is nevertheless compatible with a qualified version of realism. We will argue that, on examination, the "new ontology" is incompatible with the realism-claim implicit within the "direct realism" of contemporary analytic philosophy.

A2 Addressing the explanatory gap in representation-averse cognitive science

Figure A1 contrasts representation-friendly and representation-averse theories of perception.
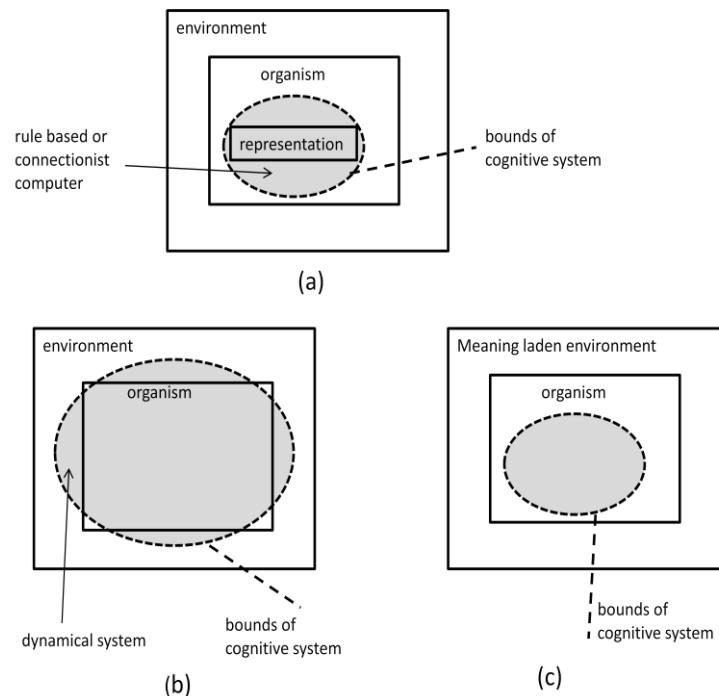
Figure A1

In figure A1(a) the organism constructs an internal representation of its environment. Such

representations serve as the basis for "surrogate reasoning"[197] about an organism-independent

environment, and the cognitive system can be[198] contained entirely within the organism. This is

the situation according to both orthodox and mainstream non-radical embodied cognitive

science. Marr's 2½-D sketch (1982) and Churchland's multidimensional face-space (2002) are

exemplars of such representations. Figures A1(b) and A1(c) relate to representation-averse

accounts of perception which address the explanatory gap, which is created by the rejection of

representations, by positing, respectively, extended cognition and a meaning-laden world.

---

[197] See Section 4.2.

[198] This is not to say that they are necessarily so contained. For example Clark espouses both action-oriented representations (Clark, 1997) and a version of extended cognition. (Clark & Chalmers, 1998)

The example of representations in multidimensional face-space (Churchland, 2002) shows

clearly how representations provide the basis for surrogate reasoning about the environment and,

importantly, the extent of the explanatory gap that must to be addressed by any representation-

averse theory of perception.

In Section 4.4 we supposed that a simplified face-recognition connectionist network was built

into a primitive organism inhabiting an environment in which type B faces were friends and type

E faces were foes. We saw that the organism was able to recognize friendly faces, and initiate

appropriate friendly behaviour, by virtue of surrogate reasoning on representations in an abstract

multidimensional face-space. And here "surrogate reasoning" could be precisely defined: that is

drawing inferences with respect to
appropriate behaviour on the basis of the
distance between points in an abstract
multidimensional face-space.[199] See figure
A2.
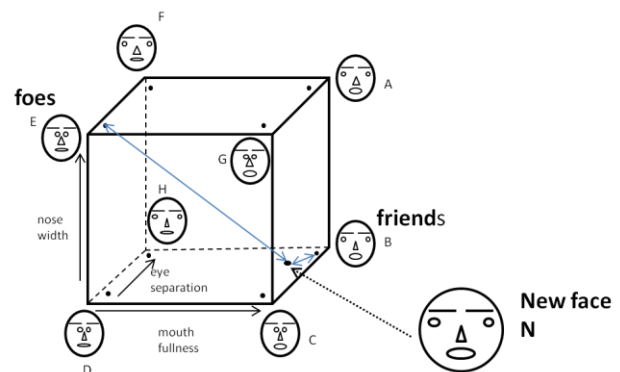
Figure A2

This highly simplified example serves to
make two key points that apply to all representation-friendly explanations of perceptually guided

behaviour. Firstly, since a representation of the environment is instantiated within the organism,

the organism has the internal resources with which to make inferences from the state of the

environment, in this case the presence of a type N face, to the appropriate behaviour. Thus

perceptually guided behaviour is explained in terms of a cognitive system, in this case a

connectionist network, which is contained entirely within the organism.

---

[199] If N is nearer, in terms of Euclidean distance in face-space, to B than it is to E
(that is: if $((x_n\text{-}x_b)^2 + (y_n\text{-}y_b)^2 + (z_n\text{-}z_b)^2) < ((x_n\text{-}x_e)^2 + (y_n\text{-}y_e)^2 + (z_n\text{-}z_e))^2$), then initiate friendly behaviour.

Secondly, the appropriate, organism-specific, behaviour is inferred from a representation of an organism-independent environment in conjunction with further, organism-specific, information contained within the organism. It is on the basis of this organism-specific information that the presence of the same type N face within the environment results in friendly behaviour in one type of organism but unfriendly behaviour in another. Thus, perceptually guided organism-specific behaviour is explained in terms of an organism-independent environment. Importantly, the environment is not inherently meaningful. The meaning of a particular environment for a particular organism, say that friendly of unfriendly behaviour is appropriate, is inferred by the organism and not simply picked up from the environment.

Such representation-friendly accounts of the perceptual process sit easily with the realism-claim. Perceptually guided behaviour is explained in terms of inferences on representations of an organism-independent world. The representations may be objective, eg Marr's 2½-D sketch (1982), or "action-oriented" with respect to the needs and capabilities of a particular organism, eg Clark (1997). However there is no difficulty in maintaining that there is an organism-independent world, even though a particular organism's perception of that world is dependent on the sense organs with which the particular organism is endowed.

On the other hand, according to representation-averse cognitive science, there is no representation of the environment within the organism. Consequently, if the organism's cognitive system were entirely contained within the organism, the organism would have no basis for reasoning about the environment and inferring the appropriate organism-specific perceptually guided behaviour; neither the environment itself nor a surrogate thereof is available to the cognitive system. Although there are significant differences in the various representations posited by mainstream cognitive science, they all provide the basis for essentially the same explanation of perceptually guided behaviour, in terms of surrogate reasoning. However, no such

explanation is available to representation-averse cognitive science. Here then is the explanatory gap.

Representation-averse cognitive science must deny that the rejection of representations creates such an explanatory gap. Indeed, the challenge to the completeness of representation-averse theories of perception has been an ongoing theme in the debate; we noted such criticisms of Gibson's theories in Section 3.1.2. As Oytam and Neilson put it, "Cognitive approaches see the theory of direct perception as incomplete. Ecological realists, in return, seem to interpret this criticism as a sign of *ontological dualism*." (2007, p. 281) As we will see, the reference to dualism in the Gibsonian rejoinder is highly significant.

In response to the charge of being incomplete, radical embodied cognitive science has two options: to deny the necessity of inferences, or to deny that the cognitive system is contained entirely within the organism. On the basis of the first denial, it can be argued that since perceptually guided behaviour is not based on inferences it does not involve the cognitive system in reasoning, surrogate or otherwise, thus there is no need to posit representations as the basis for surrogate reasoning. On the basis of the second denial, it can be argued that since the perceptual system extends into the environment beyond the organism, there is no requirement for a representation of the environment within the organism as the basis for surrogate reasoning, since the cognitive system has access to the environment proper and no surrogate is required.

As we have seen, both denials are to be found in Gibson's seminal work through to contemporary radical embodied cognitive science (RECS). We now focus on RECS as providing, currently, the most highly developed non-representational account of perception.

Along with embodiment and the rejection of representations, the two denials, that perception involves inferences and that the cognitive system is wholly contained within the organism, are essential features of RECS. We take it that proponents of RECS would agree. However our claim

that the ontological implications of the denials, a meaning-laden world and extended cognition respectively, are incompatible with the realism-claim, is contentious. As we saw in Section 3.3.3, whether or not representation-averse cognitive science is compatible with "realism", variously defined, is a matter of dispute and much depends on how that notoriously slippery term is defined. However, we argue that RECS is incompatible with the realism-claim made by contemporary direct realist analytic philosophy.

The anti-realist implications of the two denials are examined in the following sections. We maintain only that RECS must make one or the other of the denials. If RECS must make both, as some of its proponents do, our argument stands; indeed our case is the stronger.

A3 From a meaning-laden world to anti-realism

The first of the two options that RECS might take to address the explanatory gap is to deny that perceptually guided behaviour is based on inferences, and thus deny that it need involve the cognitive system in reasoning, surrogate or otherwise. If, as Gibson maintained and proponents of RECS agree, contrary to the orthodox view, perception should be understood as a direct non-inferential process, in which a sufficiency of information is picked up by an organism actively exploring its environment, then perceptually guided behaviour can be explained without recourse to representations and surrogate reasoning.

From the rejection of representations to a meaning-laden world

The further connection, between (i) the rejection of representations and surrogate reasoning and (ii) the pick-up of information from a meaning-laden world, is clear in Gibson's notion of directly perceivable affordances (1979, p. 140)[200] through to present day RECS. Chemero, a

---

[200] "The theory of affordances is a radical departure from existing theories of value and meaning. … The perceiving of an affordance is not a process of perceiving a value-free object to which a meaning is somehow added in a way that no one has been able to agree upon; it is a process of perceiving a value-rich ecological object." (Gibson, 1979, p. 140)

leading proponent of RECS, makes explicit the connection between direct non-inferential

perception, the notion of a meaning-laden world and the ontological implications thereof:

> In direct theories of perception …meaning is in the environment and perception does not depend on, meaning conferring inferences. Instead the animal simply gathers information from a meaning-laden environment." (Chemero, 2009, p. 135)

> Radical embodied cognitive science requires a new ontology, one that is at odds with today's physicalist, reductionist consensus that says that the world just is the physical world, full stop. Without a coherent understanding of what the world is like, such that it can contain meanings and is not merely physical, direct perception is simply indefensible. Thus, like earlier theories that take perception to be direct (e.g., James 1912/1976; Heidegger 1927), Gibson's ecological psychology (Gibson 1966, 1979) includes an ontology, his theory of affordances. (p. 136)

Moreover Chemero recognizes that such an ontology presents RECS with a "heavy theoretical

burden … so severe that it might outweigh all the advantages of conceiving of perception as

direct." (2009, p. 135)

From a meaning-laden world to the rejection of the realism-claim

Here then is the "ontological funny business" Chemero (2009, p. 135): RECS requires a new

ontology, an ontology in accordance with which the world "can contain meanings". Moreover,

since meanings imply the existence of organisms for which the world is meaningful, a meaning-

laden world, that is a world in which meaning is part of the fundamental ontology, would appear

to be at odds with the realism-claim. Indeed, under the heading "Do Affordances Exist Without

Animals?", Chemero notes "that it is not at all obvious at the outset that ecological psychology is

not a form of idealism…" (2009, p. 149) Although he rejects this suggestion he goes on:

> But affordances do depend on the existence of some animal that could perceive them, if the right conditions were met. Because, affordances, the primary perceivables according to ecological psychology depend in this way on animals, the ontology of ecological psychology is not a simple form of realism. It is a form of realism about the world as it is perceived and experienced – affordances, which are inherently meaningful, are in the world, and are not merely projected on it by animals. (Chemero, 2009, p. 150)

If RECS requires a new ontology according to which the "primary perceivables", that is

affordances, are both "in the world" and are "inherently meaningful" then RECS is at odds with

the realism-claim. RECS may be compatible with a "form of realism" but not the realism-claim.[201]

Importantly, it is not the notion of affordances as such which challenges the realism-claim, but rather the fact that RECS puts inherently meaningful affordances into the world, that is espouses a meaning-laden world. Non-radical embodied cognitive science can, and often does, espouse the notion of affordances, but without being at odds with the realism-claim. For example, Clark, in his influential monograph *Being There,* talks approvingly of "the Gibsonian notion of affordances" as being the "direct inspiration of the idea of action-oriented internal representations".[202] (1997, p. 172) Thus in Clark's account there is no need to suppose that affordances, "opportunit(ies) for use or interaction which some object or state of affairs presents to a certain kind of agent" are inherently meaningful elements of the world beyond the skin of the organism.

Although Clark rejects the idea of representations "as action neutral encodings of external states of affairs", he accepts the "idea of internal representations as internal states, structures or processes whose adaptive role is to carry specific types of information for use by other neural and action guiding systems." Here there is no explanatory gap that must be filled by positing a meaning-laden world; perceptually guided behaviour can be explained in terms of meaningful action-oriented representations. In contrast to Chemero, Clark can accept that affordances are "merely projected on [the world] by animals." There is no requirement for "a new ontology" according to which the world beyond the skin of the organism "can contain meanings"

---

[201] The reference to realism about the world "as it is perceived and experienced" is suggestive of the radical empiricism of James. Indeed, Chemero argues that "ecological psychology is a direct descendent of the work of James and Dewey." (Chemero, 2009, p. 30)

[202] Clark is a defender of non-radical against radical cognitive science and suggests that Gibson's "seeming to attack the notion of complex mediating inner states tout court" was a "rhetorical slip". (Clark, 1997, p. 50) He takes Gibson to be arguing against "action-neutral" representations rather than representations per se, the suggestion being that Gibson would have approved of Clark's action-oriented representations.

(Chemero, 2009, p. 136); organism-specific meanings are conferred, within the cognitive system, on an inherently meaningless world. The need for "ontological funny business" (p. 135) arises not from embodiment or affordances per se, but from the rejection of representations and surrogate reasoning thereon, and the consequent necessity of attributing meaning to the world; this is the radical move which is peculiar to radical, but not non-radical, embodied cognitive science. In gloss, RECS puts affordances in the world and meaning in the affordances.

(It is important to note that there are other senses in which meaning might be taken to be in the world, which have no such anti-realist consequences. For example, consider the notion of semantic externalism. (Lau & Deutsch, 2013) In this context, the much discussed "Twin Earth" thought experiment has been taken to show that the term "water" has different meanings if, though superficially identical in all respects, that particular substance has a different chemical composition in the two worlds. In this sense, the meaning of "water" might be said to be in the world. However the sense in which RECS requires that the meaning of, say, "water" be in the world involves a much more radical departure from mainstream thinking; unsurprisingly since RECS is avowedly radical. Following Gibson, RECS requires that "affordances are in the world, and not merely projected on it by animals". (Chemero, 2009, p. 150) And the drinkability of water might be taken to be one such affordance. Thus, the drinkability of water is simply gathered "from a meaning-laden environment", rather than inferred by means of "meaning conferring inferences" within the animal (p. 135).)

Although we maintain that Chemero's explication of the ontological implications of RECS supports our argument that RECS is incompatible with direct realism's realism-claim, it would be wrong to claim him as an ally. After all he does maintain that RECS is compatible with a form of realism, albeit "not a simple form of realism" (p. 150). However we argue below that his defence of a realism of sorts, a retreat from the antirealist position taken in his early work

(1998), does not address the issues that are peculiar to representation-averse theories of perception and does not undermine our position.

Chemero addresses the realism issues at length in his chapter on "The Metaphysics of Radical Embodiment" (2009, p. 183). However he does not address the particular issues that arise from *radical* embodiment's rejection of representations, but rather the issues common to *radical and non-radical* embodiment. Indeed Chemero uses the locution "(radical) embodied cognitive science" throughout the bulk of the chapter to show that no distinction is intended. Inevitably then, he does not address the issues relating to a mean-laden world, but rather issues relating to the organism-dependence of the senses.

However to do justice to the force of the argument from a meaning-laden world to the rejection of the realism-claim, it must be clearly distinguished from arguments which draw on the organism-dependence of sensory systems. The latter arguments claim that since organisms with different sensory systems perceive the world differently none of these perceptions can be said to be objective or to correspond to the structure of the world. Such arguments may well be valid, but they do not depend on the rejection of representations; moreover as discussed in Section 3.3 such arguments are objections to the directness-claim rather than the realism-claim. In contrast, our argument from a meaning-laden world makes no reference to sensory systems or differences between organisms, and is entirely dependent on RECS's rejection of representations. Of the two arguments, Chemero is clearly concerned with the former and does not appear to address the latter. As he sees it:

> The source of the problem with realism for radical embodied cognitive science, radical or not, is that animals perceive affordances, and affordances are animal dependent. (Chemero, 2009, p. 193)

Chemero is right to claim that this is the problem common to radical and non-radical cognitive science, but it is not the problem peculiar to radical embodied cognitive science (RECS). The

meaning-related realism-problem arises not from embodiment but from the rejection of representations.

In answer to the "problem with realism" as Chemero sees it, that is the problem common to radical and non-radical embodiment, he defends his claim that RECS is compatible with a "form of realism" by arguing for version of Hacking's "entity realism"[203] (Chemero, 2009, p. 193). However, in the terms of our distinction between the realism-claim and directness-claim, he is defending "realism" against objections which assume the former and are attacks on the latter. Chemero's version of entity realism is a response to arguments along the familiar lines that:

> Neither humans nor beetles have action-oriented representations which represent the animal-independent world exactly correctly. …. Although we may be justified in believing that there is an animal-independent external world, we have no justification to believe that our perceptions, thoughts and theories are accurate reflections of it." (Chemero, 2009, p. 187 & 188)

Here it is not the existence of an "animal-independent world" that is at issue, this is assumed, but rather the nature of our perceptions of that world. The argument from the organism-dependence of the senses is an argument against the directness-claim not the realism-claim. Arguments based on differences in perceptual apparatus can have no bearing on the realism-claim; it could be maintained *that the physical world has an existence that is not in any way dependent upon it being cognized,* even if all perception thereof is organism dependent to the maximal degree.

Chemero does not address the particular "problem with realism" which arises from the ontological implications of RECS's rejection of representations. He agrees that the rejection of representations implies a meaning-laden world, but he does not address our argument that a meaning-laden world is incompatible with the realism-claim. Our argument stands.

---

[203] See Section 3.3.3.

A4 From extended cognition to anti-realism

The second of the two options that RECS might take to address the explanatory gap is to deny that the cognitive system is contained within the organism, that is to espouse extended cognition in the strong constitutive sense. If the cognitive system extends out into the environment beyond the organism, there is no requirement for a representation of the environment within the organism, since the cognitive system has access to (indeed is partly constituted by) the environment proper and no surrogate is required.

From RECS to extended cognition

The connection between the rejection of representations and a cognitive system that extends into the environment beyond the organism, which is suggested in Gibson's analogy of the senses as "tentacles or feelers" (1968, p. 5) and his notion of direct information pickup (1979, p. 238), is made explicit in RECS. As Chemero puts it, "Radical embodied cognitive science is a form of extended cognitive science." (2009, p. 31) And Hutto and Mylin, in *Radicalizing Enactivism,* draw the connection with a particularly robust understanding of extended cognition which they label "extensive mind":

> REC ["Radical Enactive (or Embodied) Cognition"] conceives of the mind as extensive, not merely as sometimes extended. It rejects what we dub the Default Internal Mind assumption (DIM). According to the Default Internal Mind assumption, basic minds are fundamentally brain bound and ... are extended only in exceptional cases, such as cases where non bodily add-ons are required to make the achievement of certain cognitive tasks possible. (Hutto & Myin, 2012, p. 137)

The connection between the rejection of representations and the espousal of extended cognition in this strong constitutive sense is brought out by a comparison of the cognitive systems posited by the proponents of representation-friendly and representation-averse cognitive science. According to representation-friendly cognitive science, the cognitive system could be instantiated in a rule-based symbol-system or connectionist network contained wholly within the organism, but which incorporates representations of the environment. (In what follows we will

use the example of a connectionist network since this is now the focus of attention.) The

cognitive system is embrained or embodied. See figure A1(a). Moreover the hardware,

importantly but not exclusively a network of neurons, is special-purpose hardware in the sense

that it has computing capabilities that are not normally found in the environment; just as the

hardware in a lap top has computing capabilities not to be found in the desk on which it rests.

Consequently it is possible to delineate the physical boundary between the special purpose

hardware of the cognitive system and the environment.

However, proponents of RECS reject the notion of an embrained or embodied computer

processing representations of the environment, in favour of an understanding of cognition in

terms of a dynamical system comprising both organism and environment in one non

decomposable whole. See figure A1(b). As Chemero puts it:

> Agents and environments are modeled as non-linearly coupled dynamical systems.
> Because the agent and environment are non-linearly coupled, they form a unified non
> decomposable system, which is to say they form a system whose behaviour cannot be
> modeled, even approximately, as a set of separate parts. (Chemero, 2009, p. 31)

On the dynamical systems understanding of extended cognition, no in-principle distinction is

drawn between cognitively-capable computing hardware in the organism and cognitively-

incapable non-computing hardware in the environment; all hardware is potentially computing

hardware. See Section 3.2.2. Moreover the hardware of the cognitive system is a dynamical

system which extends beyond the organism into the environment. This is a bold claim: the

environment no less than the organism is literally part of the cognitive system.[204] It goes well

beyond philosophical externalism (Lau & Deutsch, 2013), "Otto's notebook" (Clark &

Chalmers, 1998), or the recognition that cognition can only be understood in terms of an

organism immersed in a particular environment, and in the human case a particular culture.

---

[204] See (Froese, et al., 2013) (Forceno, 2010) for a defence of arguments from dynamical systems to the extended mind, which stand against the "constitution v causation" objection which undermines more weakly motivated arguments.

However the term "extended cognition" is often used in a much weaker sense which, as we saw above, Hutto and Myin dub the "Default Internal Mind" assumption. For example, in the paper introducing a recent issue of *Philosophical Psychology* devoted to "the thesis of extended cognition", Arnau et al note that:

> The recent radical turn in embodied cognition, which is chiefly a form of antirepresentationalism based on a Gibsonian metaphysics of direct perception and affordances, and drawing on dynamical systems theory, has an unclear relationship with the EC [Extended Cognition] approach. It is beyond the scope of this issue to discuss these details further. (2014, p. 15)

It is, of course, exactly this form of Gibsonian "antirepresentationalism" with which we are concerned, and which we claim necessitates extended cognition in the strong constitutive sense. Nevertheless, Arnau's recent contribution is useful in making clear that much of the current discussion of "extended cognition" is not motivated by the rejection of representations and does not address the metaphysical issues raised by Gibsonian cognitive science, notably issues related to dynamical systems and affordances. Since most cognitive science is firmly wedded to the notion of representation and is Helmholtzian rather than Gibsonian, the focus of this recent work is hardly surprising. However such work is not of direct relevance to the main thrust of our argument. Here we are concerned with "extended cognition" in the strong constitutive sense, that is, as necessitated by the rejection of representations.   Moreover, it is only "extended cognition" in this sense which has anti-realist implications.

From extended cognition to the dissolution of the subject/object dichotomy

As we have noted, the notion of a cognitive system which extends beyond the skin of the organism is suggested in Gibson's seminal work. Moreover the alliance of dynamical systems and extended cognition has the resources to flesh-out key notions, notably direct information pickup and resonance between organism and environment, that were undeveloped in his

writings. But extended cognition comes at the cost of losing the separability of the cognitive

system and the environment, that is the perceiver and the perceived.[205]

We can think of the embrained-connectionist-network and extended-dynamical-system

understandings of cognition as, respectively, one and two steps removed from Cartesian dualism.

Both reject the first Cartesian dichotomy between mind (thinking substance) and matter

(extended substance), though the dynamical systems understanding goes further to reject a clear-

cut distinction between cognitively capable and cognitively incapable hardware.

However, on the dynamical systems understanding of extended cognition, the second Cartesian

dichotomy between subject and object is also to be rejected. The organism and environment

constitute one cognitively-capable non-decomposable dynamical system. See figure A1(b).

Although, from the physiological perspective, the limits of the organism could still be defined in

terms of, say, the skin, the cognitive system is no longer bounded in this way and literally

extends into the environment beyond the organism. The demarcation between subject and object,

perceiver and perceived, has been lost.

This Heideggerian dissolution of the subject/object dichotomy should not be unexpected. We

have already noted parallels that have been drawn between Gibsonian cognitive science and the

Phenomenological Tradition. Indeed Wheeler (2005) considers the Heideggerian connection in

detail and draws a distinction between "orthodox cognitive science" which operates within "the

generically Cartesian philosophical framework", and "embodied-embedded cognitive science"

which "betrays the largely tacit but robustly Heideggerian understanding of mind and

intelligence." (2005, p. 284) Moreover this Heideggerian connection is widely recognized. In

the engineering orientated journal *Presence: Teleoperators and Virtual Environments* we find:

---

[205] Interestingly, although Chemero states that "Radical embodied cognitive science is a form of extended cognitive science" (Chemero, 2009, p. 31), he has relatively little to say about the metaphysical implications of extended cognition. He makes only one brief reference in the chapter on "The Metaphysics of Radical Embodiment." (p. 202)

> Even though Heidegger's metaphysical views may seem obscure, especially in the light of most modern theories of cognitive and perceptual psychology, they are in fact not. The perceptual theory of J. J. Gibson shares a great deal with the Heideggerian orientation. The relationship is so great that it has recently been argued that Heidegger's explicitly stated ontology may in fact be used to fill in certain ontological gaps in Gibson's perceptual theory (Kadar and Effken, 1994). (Zahorik & Jenison, 1998)

From the dissolution of the subject/object dichotomy to the rejection of the realism-claim

So far we have argued that extended cognition, in the strong constitutive sense whereby the cognitive system literally extends beyond the organism into the environment, implies the dissolution of the Cartesian subject/object, self-world dichotomy. Here we take the further step to argue that such a dissolution is incompatible with direct realism's realism-claim.

To understand the force of our argument it is essential to distinguish between the two Cartesian dichotomies[206], mind v matter and subject v object, and to emphasize that extended cognition is incompatible with the latter. If extended cognition merely stood in the way of Descartes' substance dualism, then the realism-claim would not be at risk and there would be no conflict with contemporary analytic philosophy. However, as Wheeler notes, the "duality of the mental and physical" is not the only duality to be found in Descartes:

> There is also the duality of the subject and object, that is the dichotomy between the individual cognizing subject and the world of objects about which the subject has beliefs. (2005, p. 22)

Wheeler goes on to reject this subject/object dichotomy in the name of "Heideggerian cognitive science" (2005, p. 285). However such a move is not open to the direct realist who must defend direct realism's realism-claim which is grounded on the distinction between "perceiving or

---

[206] As Wheeler sees it: "That these two Cartesian dualities are conceptually separable is apt to be obscured by the fact that Descartes tends to position the epistemic anterface between the subject and object at his proposed metaphysical interface between the mental and the physical with the mind lining up with the subject, and the physical world (including the body) lining up with the world of objects." (Wheeler, 2005, p. 23)

thinking subjects" and "elements of this physical world".[207] Thus, although direct realists might well reject mind/matter dualism, they must retain subject/object dualism. Indeed Cartesian subject/object dualism is an unspoken assumption within both mainstream cognitive science and much of contemporary analytic philosophy.[208] Wheeler expresses well the hold that this assumption maintains:

> At first sight it might seem that drawing attention to the assumed primacy of the subject-object dichotomy in this way is a sorry waste of everyone's time. After all, what could be more uncontroversial, or matter of philosophically uncluttered commonsense, than the fact that each of us can make a distinction between, on the one hand, his or her own self , and on the other, a world of objects which exists independently of us …" (Wheeler, 2005, p. 23)

Whether or not this is "uncluttered commonsense", is open to dispute and Wheeler takes issue from the Heideggerian perspective. However, the "subject-object dichotomy" is clearly essential to the spirit of, and the worldview implicit in, the direct realism espoused by contemporary analytic philosophers. Indeed Brewer talks in terms of the "mind-independent physical objects we all know and love" (2007, p. 87) and a "commonsense realist worldview" (2004, p. 61). More explicitly, in his recent defence of direct realism, Searle endorses "the account that bottoms out in what I think of as the real world that exists in a way that is totally observer independent and ontologically objective." (Searle, 2015, p. 223) [209] [210]

---

[207] In full: "The physical world has an existence that is not in any way dependent upon it being "cognized". ….the purely physical elements of this physical world of ours would, or at least could, be just as they are even if no perceiving or thinking subjects had ever come into existence." (Smith, 2002, p. 2)

[208] Though questioned within the Phenomenological Tradition, as in Husserl, Merleau-Ponty and Heidegger.

[209] Similarly, Searle in defending his scientifically sophisticated common sense introduces five "default positions" which he believes to be true and not open to legitimate attack (Searle, 1999, 10). The first, "There is a real world that exist independently of us ….", is clearly grounded on the subject object dichotomy.

[210] Moreover, from the other side of the debate, Varela et al (1991, p. 9) take the assumption that "there is a separate subjective we" to be a mark of the "realism or subject/objectivism" that they reject.

Moreover, the subject/object dichotomy is not only essential to the letter and spirit of direct

realism (metaphysical or epistemological), but an assumption on which the debates within the

recent philosophical literature are grounded. For example, in introducing his discussion of "the

nature of perception" Armstrong notes: "To say of any organism that it can perceive is to imply

that it can acquire a certain amount of knowledge of the world around it by means of sense

organs called the senses." (1961, p. 105) Here we have a definition of perception which would

be acceptable to both direct and indirect realists, and which sets the scene for the debate between

them. Clearly implicit, is the notion of a perceiving organism, an "it", separate from the "world

around it" with its "sense organs" at the organism/world interface.

More recently Gendler and Hawthorne introduce a collection of papers under the title *Perceptual*

Experience (2006) thus:

> Much contemporary discussion of perceptual experience can be traced to two
> observations. The first is that perception seems to put us in direct contact with the world
> around us: when perception is successful, we come to recognize – immediately – that
> certain object have certain properties. The second is that perceptual experience may fail
> to provide such knowledge: when we fall prey to illusion or hallucination, the way things
> appear may differ radically from the way things actually are. (Gendler & Hawthorne,
> 2006, p. 1)

Again we can see that the discussion is grounded on the distinction between perceiver and

perceived, that is an "us" and a "world around us". This distinction is not at issue. What is at

issue is the possibility of reconciling "direct contact" with the facts of illusion or hallucination.

In our terms, it is only the "directness-claim" that is at issue.

Without this distinction between an "us" and "world around us" it would be difficult to make

sense of recent debates within the philosophy of perception. Consider the much discussed

Phenomenal Principle which plays a key role in the debate between disjunctivist and

representationalist direct realism:

> If there sensibly appears to a subject to be something which possesses a particular sensible quality then there is something of which the subject is aware which does possess that sensible quality. (Robinson, 1994, 32)

Clearly implicit in the Phenomenal Principle, is the notion of a "subject" and a world of "something[s]" distinct from the subject. Moreover, when the Principle is questioned, it is not the subject/object dichotomy which is in dispute, only the nature of that "of which the subject is aware". The subject/object dichotomy provides the conceptual framework within which the debate is conducted, and is assumed by disjunctivists and representationalists alike.

Our conclusions with respect to the anti-realist implications of representation-averse cognitive science are inevitably more tentative than those of Chapter 4. What constitutes the realism component essential to direct realism, and what would count against it, will inevitably remain points of contention. As Williamson puts it in the *Oxford Companion to Philosophy*:

> Realism and anti-realism: Primarily directions rather than positions. To assert that something is somehow mind-independent is to move in a realist direction; to deny it is to move in the opposite direction. (1995, p. 746)

The question is how far in the "realist direction" does the direct realist need to hold the line. Although it would be difficult to legislate on just how "realist" a "direct realist" needs to be to remain true to the cause, a Heideggerian dissolution of the perceiver/perceived dichotomy, no less than a Merleau-Pontian meaning laden world, is clearly a retreat too far in the anti-realist direction. Thus we conclude that RECS is at odds with the "realism" of both mainstream cognitive science and contemporary direct realist analytic philosophy. Gibsonian representation-averse cognitive science, as exemplified by RECS, does not constitute a counter example to our central thesis.

Tellingly, the final chapter of Wheeler's *Reconstructing the Cognitive World*, in which he rejects the subject/object dichotomy and welcomes what he sees as a Heideggerian turn in cognitive

science, is titled "A Re-Beginning: It's Cognitive Science, But Not as We Know it." (2005, p. 283)

A5 Some provisional conclusions and caveats

As we have noted at a number of points in the discussion, our arguments with respect to representation-averse cognitive science and the realism-claim are more tentative than those in the main body of the dissertation. Here we present provisional conclusions along with caveats thereon.

1. The realism-claim is a claim about how the world is, irrespective of whether, and how so, it is perceived. Thus considerations with respect to the process of perception, in our or other species, though telling with respect to the directness-claim can have no bearing on the realism-claim. For this reason the argumentation is markedly different in form and necessarily less direct: it is argued that representation-averse cognitive science cannot explain perceptually guided behaviour without positing an ontology which is at odds with the realism-claim.

2. Although proponents of representation-averse cognitive science tend to agree that it requires "a new ontology" (Chemero, 2009, p. 136), the second step in the argument, from a new ontology to the rejection of the realism-claim, is more contentious. As we have seen, some are explicitly anti-realist (Varela, et al., 1991); indeed some are sympathetic to Berkeleian idealism (Oytam & Neilson, 2007). But, on the other hand Chemero, although a onetime anti-realist (1998), now argues that representation-averse theories are compatible with a realism of sorts (2009). And although Wheeler (2005) talks approvingly of a Heideggerian cognitive science he does not endorse the anti-realist claims.

3. We have proposed two lines of argument to show that the "new ontology" of representation-averse cognitive science is incompatible with the realism-claim: from a "meaning-laden world", and from "extended cognition". However, both notions are often construed in a sense which has no anti-realist implications. Thus the burden of our argument rests on establishing that, in the context of (avowedly radical) representation-averse cognitive science, these notions must be understood in a strong constitutive sense which does have such ontological implications. We have made a prima facie case, but recognize that legitimate concerns remain.

4. It could be objected that the inference from a "meaning-laden world" to anti-realism is a bad one, since things can instantiate properties that are defined in terms of subjects even if no such subjects actually exist; thus the world may be both meaning-laden and subject-independent. And, moreover, even if some properties aren't entirely objective (the "meaning-laden" ones), that does not mean that all of reality is subjective. To address the first point, further work is needed in order to demonstrate more conclusively the connection between representation-averse cognitive science and "a meaning-laden world" construed in the strong constitutive sense which allows that inherently meaningful Gibsonian affordances are constituents of the world and not merely projected on it. The connection between radical embodied cognitive science and the Phenomenological Tradition needs further investigation; the parallels between the work of Gibson and Merleau-Ponty, eg in Sanders (1993) (1997), are particularly interesting. On the second point, further exegesis is needed to demonstrate more conclusively that contemporary direct realists are in fact defending the stronger "realism-claim" as defined in Section 1.2.1.

5. Similarly, the argument from "extended cognition" can be challenged on the grounds that, the subject/object distinction is not compromised by the claim that, for example, the cognitive system is non-decomposable in the sense that its behaviour cannot be understood in terms of the distinct behaviours of its parts. To address this type of objection, further work is required to demonstrate more conclusively the essential connection between representation-averse cognitive science and "extended cognition" construed in the strong constitutive sense according to which the cognitive agent is not merely explicable in terms of, but literally is, a dynamical system which extends into the environment. Here there may be fruitful connections between the notion of a Heideggerian turn in cognitive science (Wheeler, 2005) and new ideas in computation, for example Wolfram's (controversial) principles of "computational reducibility" and "computational equivalence" (Wolfram, 2002).

[Final Draft 9.docx
84,820 words including footnotes, references and Appendix A]