Perceptual constancy

by

Alessandra Buccella

Bachelor's degree in Philosophy, Università degli Studi di Milano, 2011

Master's degree in Analytic Philosophy, Universitat de Barcelona, 2013

Master's degree in Philosophical Sciences, Università degli Studi di Milano, 2013

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Alessandra Buccella

It was defended on

April 6, 2020

and approved by

Mazviita Chirimuuta, Associate Professor, History and Philosophy of Science, University of Pittsburgh

Anil Gupta, Alan Ross Anderson Distinguished Professor of Philosophy, University of Pittsburgh

Robert Brandom, Distinguished Professor of Philosophy, University of Pittsburgh

Wayne Wu, Associate Professor, Philosophy, Carnegie Mellon University

Dissertation Director: Mazviita Chirimuuta, Associate Professor, History and Philosophy of Science, University of Pittsburgh

Perceptual constancy

Alessandra Buccella, PhD

University of Pittsburgh, 2020

We perceive objects and events in a way that makes it possible to act, react, think, believe, etc. in reliable and predictable ways. To explain this perceptual stability, as well as its behavioral consequences, theorists invoke a set of capacities known as *perceptual constancies.* Thanks to constancies, perceivers latch onto what's unchanging in the world even though sensory stimulation is in continuous flux. In this dissertation, I present and defend a new view of both perceptual constancy and perceptual objectivity, i.e. the capacity of perception to present the world as mind-independent. According to the traditional view, perceptual constancy is the capacity of perceptual systems to recover perceiver-independent properties of distal objects from a largely ambiguous proximal stimulus, 'discounting' contextual, perceiver-dependent information. I argue that the traditional view should be rejected because it is, on the one hand, too 'visuo-centric', and, on the other hand, unable to fully explain the roles that constancy plays in our lives. These roles include guiding action and enabling the stable conscious experiences that ground our perceptual judgments. The view I favor, which I call "Relational Invariance view", holds that constancy is the capacity to track invariant relations within the perceptual scene or between some element in the scene and the perceiver. These invariant relations are specified by patterns of variation in the proximal stimulus over time, and perceivers can sometime directly control this variation through movement. This view explains the role that, intuitively, perceptual constancy plays in guiding motor action and in a wide variety of perceptual recognition tasks, where recovering perceiver-independent properties seems unnecessary. The Relational Invariance view is then tied to a new view of perceptual objectivity, whose core insight is that the 'job' of perception in enabling the

experience of a mind-independent world is not to ‘abstract away’ from any sort of perspectival or contextual influence, but rather to ‘embrace’ these influences as intrinsic to the very idea of what it means to perceive the world for creatures like us.

Table of contents

List of figures

Preface

For as long as I can remember, I wanted to communicate; to express my thoughts, feelings, and ideas clearly enough for people to listen and understand. Some people write poems, others compose music, others dance, fight, play sports; for me, it was philosophy. I ran into this quotation not long ago, while reading a collection of essays putting together the two great loves of my life: Philosophy and Basketball. The quotation is from an essay titled "Hoosiers and the Meaning of Life", by Michael E. Peterson:

> "One function of philosophy is to notice interesting features of the world so that its depth and wonder can be thoughtfully explored. Without having to go anywhere or do anything extraordinary, we simply need to see the clues around us every day."

'Depth' and 'wonder' are exactly the two words that I have been implicitly associating with the activity of philosophizing since I discovered it. The Greeks saw in human beings' capacity to wonder (θαυμάζειν) the source of philosophical inquiry, and in Dante's *Inferno* it is Ulysses who pronounces the iconic verses:

> "Considerate la vostra semenza:
> fatti non foste a viver come bruti,
> ma per seguir virtute e canoscenza."

Through these words, Dante makes of Ulysses – who in Homer's Odyssey was described as one of the most cerebral, thoughtful, and inventive heroes – a symbol of humans' burning curiosity and desire for knowledge. My early studies of classics and Italian literature had a profound influence on who I am today. For this, I want to express my gratitude one more time to Renzo Talamona, whose wisdom, curiosity, and capacity for wonder have been and continue to be an inspiration. I am proud to have been one of his last students before he retired.

At the University of Milan, Clotilde Calabi has been a mentor and a model of thoughtfulness, intelligence, wit, and kindness. She first introduced me to academic philosophy and encouraged me to pursue this career while never hiding the hard truths from me. I learned a lot from her courses, as well as from our one-on-one interactions and from the conferences and workshops she involved me in as a Master's student.

There is also a long list of people here in Pittsburgh who deserve a special thank you. My advisor Mazviita Chirimuuta and my other committee members Wayne Wu, Anil Gupta, and Bob Brandom: you taught me so much, and I will be forever grateful for the time we spent together, for your advice, and your honesty along the way. Jessica Gelber, Mark Wilson, Kate Stanton, and James Shaw also helped me in different ways at different stages of my PhD journey. My fellow graduate students Alison Springle, Tom Breed, Max Tegtmeyer, Dan Kaplan, Vivian Feldblyum, Chuck Goldhaber, and Stephen Mackereth, as well as my dear friend Ariel Kornhauser, have been my second family. It is hard to put into words how much of what I have accomplished in the last 6 years wouldn't have been possible without them. Away from Pittsburgh, Manuel García-Carpintero, Genoveva Martí, Tony Chemero, and Tom Polger deserve a special mention.

I want to thank my parents and my brother, for always making me feel their love and support across continents. I know that my career choice, the distance, and other aspects of my life as an adult woman haven't been easy for you to understand and accept, but we have always remained a family through it all: thank you. Thank you also to my partner Adelina, whose love against all odds has been my anchor through difficult times. Last but definitely not least, I want to thank my beloved grandmother Luciana. Knowing that she won't see me accomplish this breaks my heart. She was my rock, my best friend, and my hero. This work is dedicated to her. May whatever I do next make her proud.

1.0 The traditional view of constancy and its limitations

1.1 Introduction: what is perceptual constancy?

Consider the following familiar perceptual experiences. We perceive a door as maintaining a stable shape even though it projects differently shaped images on our retina as we move in respect of it (Figure 1). We also perceive objects illuminated differently at different times of the day as maintaining their colors throughout changes in illumination (Figure 2). Looking at other sense modalities, we auditorily experience a sound's timbre (e.g. Isaac, 2018) and loudness (e.g. Zahorik & Wightman, 2001) as remaining constant even when the context and the spatial relations between the perceiver and the sound source change over time. When we touch a surface, we experience its texture as remaining the same across different speeds at which our tactile receptors 'scan' the surface (Boundy-Singer, Saal, & Bensmaia, 2017). We also experience our bodies as moving smoothly and effortlessly as we walk around even though our bones, tendons, and muscles react to and produce an incredibly varied class of forces at every step. And we experience balance when standing in an upright position even though our body is actually swaying back and forth (e.g. Balasubramaniam et al., 2000; Lee & Lishman, 1975; Riley et al., 2011).

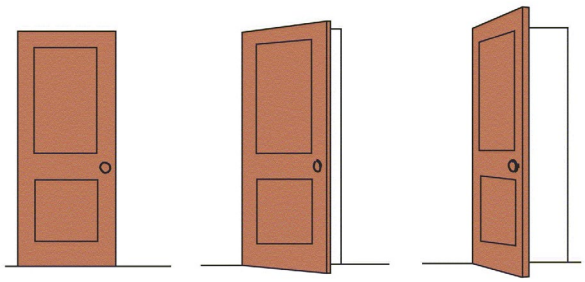

Figure 1. Standard illustration of 'shape' constancy.

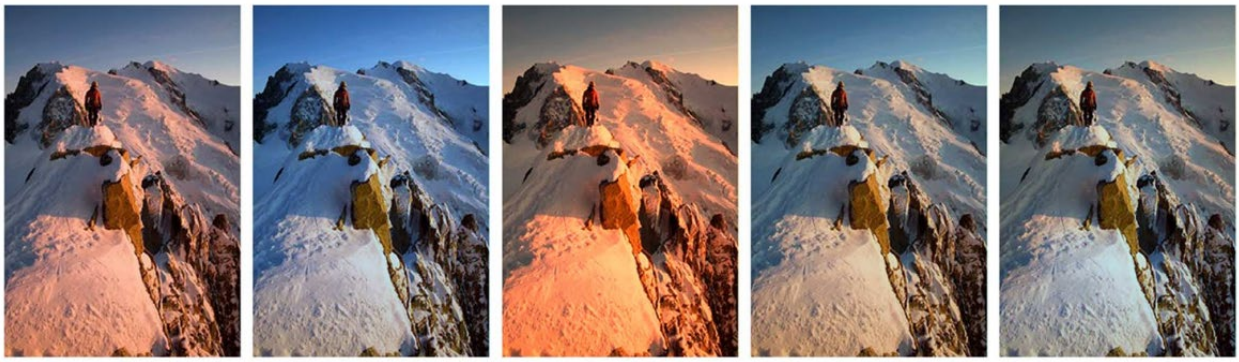

Figure 2. Standard illustration of 'color' constancy.

Despite these experiences being quite ordinary and happening very often in our everyday life, the fact that they occur at all is quite striking. Rich and stable perceptual experience, after all, depends largely on the stimulation of our sensory organs' receptors, and this stimulation itself is anything but stable. The properties of proximal sensory stimuli such as retinal images or the soundwaves hitting our eardrums are always changing due to the constant fluctuations in the levels of energy hitting sensory receptors. Yet, most of the time we perceive the external world in a way that makes it possible to act, react, think, believe, etc. in reliable and predictable ways; in a word, we are capable of perceptually experiencing *invariance* despite an enormous amount of variance in the way our sensory organs get stimulated.

But where does this experienced invariance 'come from'? What is it about our perceptual systems that makes them able to 'transform' incoming unstable and fluctuating sensory inputs into experienced stable objects, properties, or events? And *what* exactly remains invariant during these experiences? To explain this experienced invariance, as well as its behavioral consequences, theorists invoke a set of capacities known as *perceptual constancies.*

A correct understanding of the notion of perceptual constancy is key to any theory of perception that wants to explain the links between conscious experience, perceptual judgment, and

action. Indeed, the constancies seem to play a very important role in all these domains. Thus, in what follows I will not defend this claim, but rather take it as my starting assumption: whichever way we define the constancies and whichever way we describe their underlying mechanisms, without them it would be impossible to issue accurate perceptual judgments, successfully perform perceptually guided actions, or even simply have stable and coherent perceptual experiences. Therefore, I suggest that we assume this thesis as our starting point:

Perceptual Invariance: Perceptual constancies are the capacities of perceptual systems responsible for the kind of perceived invariance *necessary* for stable conscious perceptual experience, accurate perceptual judgments, and successful perceptually guided motor[1] actions.

But how do perceptual constancies work? How do these capacities of perceptual systems bring about the effects we are all familiar with? In this chapter, I start answering these questions in three main steps. First, I will present the most common way in which philosophers and scientists have been thinking about perceptual constancy, as it results directly from a more general way of understanding perception and its main goals. I will refer to this way of thinking about constancy as 'the tradition'. Second, I will isolate the tradition's theoretical assumptions which characterize and structure the approach by means of the formulation of two theses.[2] Third, I will offer some reasons why I think that the traditional approach does not tell us the whole story about perceptual constancy. In particular, even though perceptual constancy seems to quite obviously play a role in guiding action (how can you act on an object if you don't perceive it in a stable way across changes in the way it impacts your sense organs over time?), the traditional view is not in a position to account for it. The traditional approach has gone mostly unchallenged as a general framework,

[1] I specify that I am talking about 'motor' actions in order to rule out, for instance, the formation of an intention. What I call 'action' from now on, therefore, should be thought of as motor action, that is, the 'phase' of intentional action that involves perceptually guided bodily movement.

[2] Even though it might still be that single particular philosophers or scientists within the tradition would not accept both of them in the formulation I chose.

though some alternatives have been proposed with respect to particular cases, which I discuss later on. In what follows, I will make explicit the theoretical background of the traditional view and argue that such a background constrains what counts as perceptual constancy in ways that are by no means obviously correct.

1.2 The tradition

When we look at figures 1 and 2, we seem to be aware of *something* remaining constant or invariant while something *else* changes. The traditional view identifies the invariant element with the 'absolute', perceiver-independent properties of objects, which can be represented as unchanged despite contextual changes. What strikes me as surprising, however, is how this view of constancy seems to be taken for granted, as if it were a theory-neutral description of a perfectly straightforward capacity. In particular, it has been taken for granted that, in order to understand and explain perceptual constancy, we should focus on properties like shapes, colors, or sizes and the way in which perceptual systems allegedly 'retrieve' them from proximal stimuli.

For example, this is how Irvin Rock (1975) describes constancy: "Perceived qualities such as color, size, and the like tend to remain constant *despite* the fact that the proximal stimulus, for example, the retinal image of objects, *is continually changing*" (cited in Hatfield, 2009, 181; emphasis mine). Hochberg (1988; cited in Brown, 2003) defines perceptual constancy as "the constancy of perception of the fixed properties of distal objects, despite variation in the proximal stimuli from objects." (251-252). More recently, Tyler Burge (2010) discusses perceptual constancy and its role within perception more generally. He understands constancy as the perceptual capacity "systematically to represent a particular or an attribute as the same despite

significant variations in registration of proximal stimulation" (2010, 408). This group of traditional definitions has a common 'theme': perceptual constancy is the capacity of perceptual systems to represent the actual properties of objects in the world in spite of generally unstable and fluctuating proximal stimuli over time.

The traditional view claims that, because proximal stimuli are unstable and fluctuating, the 'job' of a capacity like constancy is to 'compensate for' this instability in order to reach an accurate representation of how things 'really' are.[3] Indeed, most scientists studying constancy in the traditional way implicitly assume that to display perceptual constancy is to solve so-called 'inverse problems'. Inverse problems are also known as 'underdetermination problems', because proximal stimuli alone cannot specify what kind of physical layout, among many different possibilities, is actually out there: proximal stimuli underdetermine the distal layout. For instance, various trapezoidal tables and a rectangular one can cause the same two-dimensional retinal image when viewed from certain angles, or many different combinations of light frequency and Surface Spectral Reflectance properties (SSRs) can cause the same pattern of retinal cone activation (i.e. the same visual proximal stimulus). This is a problem, because if the 'job' of perceptual systems is to represent accurately objects and properties as they really are, then these representations must be 'extracted' in multiple steps by analyzing proximal stimuli into context-dependent and context-independent causes.

According to the traditional view, perceptual constancy depends on solving the underdetermination in proximal stimuli by reverse-inferring the right distal array, so that these

[3] Advocates of this view of perceptual constancy include perception scientists in the so-called 'constructivist' camp, sympathetic to the ideas of Helmholtz (1867) and Marr (1982). The main tenet of this type of views is that, because of the ambiguity of proximal stimuli, perception of how things are independently of us must depend on a capacity to 'construct' a representation of the object and its properties which can be kept 'fixed' throughout changes in contextual and perceiver-dependent elements over time.

stimuli can be connected to the right external causes (e.g. a perceiver-independent property like a particular geometrical shape). To do so, context, idiosyncrasies of the perceiver, etc. must be first estimated and then discounted (cfr. Burge 2010; Rescorla, 2015). Here is Rescorla:

> Perception solves an underdetermination problem. *The perceptual system estimates environmental conditions, such as the shapes, sizes, colors, and locations of distal objects*. It does so based upon proximal stimulations of sensory organs. The proximal stimulations underdetermine their environmental causes. For instance, a convex object under normal lighting generates retinal stimulations ambiguous between at least two possibilities: that the object is convex and that light comes from overhead; or that the object is concave and that light comes from below…In general, then, retinal input underdetermines possible states of the distal environment.
>
> Perceptual constancies are capacities to represent properties or entities as the same despite large variation in proximal stimulation. To varying degrees, human vision displays constancies for numerous properties, including size, shape, location, colour, depth, and motion. How does the perceptual system achieve constancies? By using "implicit assumptions" to discount variations in proximal stimulation. (Rescorla, 2015, p. 2; emphasis mine).

Setting aside some slight differences,[4] the traditional view claims that the perceptual stability we all experience depends on perceptual systems successfully recovering the perceiver-independent properties of objects after discounting the rest of the information contained in ambiguous proximal stimuli. In what follows, I will identify the traditional view with the following thesis:

---

[4] For instance, Rock, (1975, 1977) following Helmholtz and his 'unconscious inference' account of how the visual system transforms proximal subjective sensations into distal object representations, holds that proximal stimuli should be considered an 'early stage' of perceptual processing. According to Rock, the role of proximal stimulation is to provide one of the two 'premises' of an unconscious inference-like procedure which has the (possibly accurate) attribution of a specific property to an object as its 'conclusion', the other premise usually being an abstract rule or principle (itself 'located' somewhere in the brain) which 'tells' the visual system how to interpret the proximal stimulus. As Gilchrist (2012) notices, the idea that perception (and perceptual constancy) should be understood as a process consisting in two separate stages (i.e. sensory stimulation and rule-based interpretation) was very much opposed by the Gestalt psychologists (Koffka, 1935; Kohler, 1947). However, criticisms to this idea have come from within the tradition, too. For example, Burge (2010) rejects the existence of (explicitly or implicitly) represented 'inferential rules' in the brain according to which we interpret proximal stimuli.

Traditional Constancy Thesis: Perceptual constancy is the capacity of perceptual systems to recover the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli.

But the tradition also accepts *Perceptual Invariance*, and, consequently, the following two theses, as well:

Constancy in Judgment: Recovering the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli is necessary for accurate perceptual judgment.

Constancy in Action: Recovering the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli is necessary for successful perceptually guided motor action.

Setting aside *Constancy in Judgment* for now, in the next section I am going to argue that *Constancy in Action* is false. This puts pressure on the traditional view as a whole, because if *Perceptual Invariance* is intuitively true, then denying *Constancy in Action* just means denying *Traditional Constancy. Constancy in Action* is false because a counterexample can be found to the idea that successful perceptually guided motor action requires the recovery of perceiver-independent properties. Therefore, if we want to maintain that successfully exercised perceptual constancy is in fact required for successful perceptually guided motor action, then perceptual constancy *cannot* be the capacity to recover perceiver-independent properties.

Before moving on, however, a clarification is needed. Let's go back to figures 1 and 2: while helpful to introduce the notion of perceptual constancy at the pre-theoretical stage, we should be careful in interpreting the relationship between pictures like those and perceptual constancy. Take this other probably familiar example. When standing in the middle of a tree-lined road, we experience a line of trees of the same size located at different distances from us, even though the farther away trees project smaller images on the retina than the closer trees. Indeed, this is what Christopher Peacocke (1983) claimed, based solely on what seemed to him like an accurate

description of his own experience: when looking at the trees without moving, the ones farther away look to be of the same size as the closer ones, even though their images on the retina are different in size.

The experience of 'Peacocke's trees' is what most non-philosophers (as well as many philosophers and psychologists) would take as a paradigmatic example of a successful display of perceptual constancy. Consistently with the traditional definition, the most natural explanation for this experience is that the visual system produces a representation of the perceiver-independent and perspective-independent size shared by all the trees. That is what we primarily perceptually represent, and perceptual constancy is the capacity to recover such a representation from the retinal image.

Figure 3. Something similar to Peacocke's trees.

However, it is a mistake to take these 'static' images as themselves examples of successful displays of perceptual constancy. Our judgments regarding the objective properties of objects in experiences like Peacocke's trees might actually be knowledge-based and inferential, rather than

purely perceptual. For instance, Granrud (2009, 2012; Granrud & Shmechel, 2006) showed that in situations analogous to Peacocke's trees, adults issue more accurate 'objective' size judgments than children. He argues that during development we learn how to employ a 'strategy' that allows for more accurate estimations, so that older perceivers are generally 'strategy-users', while younger ones aren't. The strategy in question consists in explicitly interpreting phenomenal experience through the lens of what one *knows* about how objects typically look when they are far away, namely, that they appear smaller even though they have the same size as when seen up close.

Indeed, in Granrud's experiments younger children (4-8 years), who mostly lack the capacity to explicitly reason on the basis of previously acquired knowledge, report objects seen at a distance to be smaller than they actually are. Older children (9-11 years) and adults, on the other hand, tend to (1) explicitly report that they reached their judgment by 'reasoning' about the effects of distance on size appearance, and (2) be either almost perfectly accurate or 'overconstant' (i.e. they attribute to objects a bigger size than the one they actually have). Employing the strategy has the effect of 'inflating' size judgments, arguably because strategy-users end up over-compensating for distance. Granrud's results support the idea that reports about the actual properties of experienced objects result from a cognitive inference rather than these properties being directly represented in perception. Therefore, introspection is not the best way to decide what a capacity like perceptual constancy consists in, or whether it is perceptual at all. [5] In situations like the experience of Peacocke's trees, it is indeed plausible that, knowing the effects of distance on, say, size appearance, 'mature' observers who are trying to issue an accurate size judgment will rely on explicit reasoning, i.e. what Granrud called 'the strategy'. They can compare the different

[5] Not everyone ruled out the possibility that perceptual constancy is in fact a full-blown cognitive, inferential capacity: this was, for example, Russell's (1912) view.

perspectival appearances of the trees, which are present simultaneously on the retina, and use their knowledge about distance to make the appropriate 'calculations' and arrive at a perspective-independent size estimate. This suggests that the experiences we have in these static, synchronic cases may not result directly from successful displays of perceptual constancy, but from a cognitive inference.

If Granrud is right and, at least in the case of size constancy, we in fact *learn* to cognitively compensate for distance when estimating sizes, the next question becomes: How is this ability acquired? The strategy is a way of 'correcting' experience with knowledge, but this knowledge must nonetheless come from somewhere. How does a child learn how spatial relations work, so that she can learn and employ the strategy? The answer I find most plausible is: by being frequently exposed to cases in which objects move away from or closer to her, *yet they don't look like their size is changing*. A central feature of visual experience is that it is continuous and extended in time, and as time passes the visual scene can undergo some change, in particular *spatial* change: objects change their position in space, we change our position in space, etc., and we perceptually keep track of these changes through variation in proximal stimuli. Therefore, the strategy that allows to issue accurate size judgments in *static*, *synchronic* cases like Peacocke's trees can be learned only by experiencing invariance and stability in objects' sizes *while* distance changes, that is, in cases that are *dynamic* and *diachronic*.

If all this is true, then it is fundamental to the very notion of perceptual constancy that it is displayed in dynamic situations, that is, *while* change in proximal stimuli over time occurs. Perceptual constancy is the capacity to stabilize perception of the distal world as proximal stimuli change *over time*. In other words, the static cases we started from are not themselves displays of constancy, but *depend on* displays of constancy, which in turn are phenomena occurring over time,

in which proximal stimuli change and yet we experience something as remaining invariant (whether or not this ends up being a perceiver-independent property of an object or something else). Dynamic, diachronic constancy is the fundamental notion. With this in mind, I turn to the issue of whether the traditional view can account for the role of perceptual constancy in motor action, and thus to why *Constancy in Action* is false.

1.3 Constancy in action and experimental methodology

It is widely accepted in vision science that the capacity to consciously perceive objects as having certain 'objective' properties (or the capacity to attribute such properties to objects in judgment) and the capacity to rely on perception to guide motor action can be possessed independently of one another (Milner & Goodale, 1995, 2008; Goodale & Milner, 2013; Mishkin & Ungerleider, 1982). It is also widely accepted that the two capacities are realized by two different neural pathways the cortical visual system splits into and rely on different kinds of visual representations. The capacity to categorize and recognize objects on the basis of their properties depends on the ventral stream and relies on 'allocentric' (which, for our purposes, can be equated to 'perceiver-independent') representations of distal objects. On the other hand, the capacity to coordinate movements and act on the basis of perceptual information depends on the dorsal stream and relies on 'egocentric' representations and on the capacity of the system to stabilize retinal input throughout spatial changes (due to saccadic movements of the eye or movement of external

objects) enough to appropriately guide motor action (e.g. Deubel et al., 2010; Helmholtz, 1867; Wu, 2014).[6]

A striking source of evidence for the separation of the two capacities, both functionally and anatomically, comes from Milner and Goodale's study of a woman known under the pseudonym 'Dee Fletcher' (or 'D.F.'). D.F. suffered damage to her ventral visual stream due to carbon monoxide poisoning, and as a consequence of this has a condition called 'visual object agnosia'. D.F. can visually experience features like color and texture, but is incapable of attributing such features to objects. In fact, she cannot perceive discrete distal objects at all. However, at least in familiar environments, she is capable of moving around, purposefully interacting with objects, and in general performing actions that clearly require visual guidance and quite sophisticated visuo-motor coordination (see the numerous anecdotes in Goodale & Milner, 2013).

The case of D.F. is relevant for the present discussion because of its implication for the traditional view of perceptual constancy. Indeed, we could talk of D.F. (i.e. her visual system) as, on the one hand, lacking the capacity that the tradition identifies with constancy, namely the capacity to represent perceiver-independent properties of distal objects (due to ventral stream damage), but, on the other hand, seemingly having the capacity to perform perceptually-guided motor actions, which also seems to require constancy (recall the point about stability with which I started the chapter). This in turn implies that *Constancy in Action* is false. But if we find it plausible that D.F.'s visual system is capable of constancy, which in turn depends on the broader point that constancy seems intuitively necessary for perceptually guided action, then the Traditional Constancy thesis itself is at risk.

[6] Indeed, Wu (2014, p. 385 and elsewhere) calls this capacity of the dorsal stream to stabilize sensory inputs so that egocentric relations with the world are represented and can guide action "spatial constancy".

We can show this by taking a closer look at the experimental methodology employed by perception scientists within the tradition. Indeed, the tradition has approached the empirical study of perceptual constancy mostly by relying on subjects' reports. The idea is that, when they report, subjects are simply describing their experience, or what they *take their experience to be.* However, reports can be influenced by a huge amount of different factors, the most obvious of which is certainly the way in which the experimenter formulates her *instructions* for the subjects (e.g. Arend and Reeves 1986, Wagner 2012 for a review). If both the instructions and the experimental setting (both conceived with the traditional definition of constancy in mind) prime subjects to interpret their experience in terms of discounting of contextual information and recovery of the 'actual' properties objects possess, then these experiments do not reveal much about the nature of the capacity that supposedly is being employed.

Moreover, the functional separation between perception-for-judgment and perception-for-action in the visual system discussed above suggests that a view of constancy that can explain its role in enabling judgments (i.e. reports) does not automatically explain its role in enabling perceptually guided action, at least unless the functional separation is acknowledged and more work is done. Experiments based on explicit experience reports or matching of samples (which implicitly relies on the subject making a judgment about the property in respect of which she thinks she is comparing the samples) cannot immediately be considered evidence for the correctness of the Traditional Constancy thesis. A definition of perceptual constancy has to be tested in the context of *both* the capacity to issue accurate judgments *and* the capacity to act successfully. The traditional view is importantly biased towards seeing the visual system as itself a system that judges (i.e. identifies specific properties and decides to which object or set of objects they should be attributed to).

If I am right, then we have a reason to claim that, in order to really understand what perceptual constancy is and how it works, we need to put the traditional approach aside for the moment. A good example of how worries regarding experimental methodology can lead to a rejection of the view the experiments supposedly tested can be found in the debate regarding *color constancy*. On one side, we find the traditional approach, based on the idea that color constancy with respect to a specific surface is achieved when the perceiver-independent color property of that surface is recovered through the solution of the inverse problem, hence the name 'inverse optics' (Brainard, 1998; Maloney, 1986, 2003; Maloney & Wandell, 1986). On the other side, we find an alternative suggestion, called *Relational color constancy* (Amano et al., 2005, 2006; de Almeida et al., 2010; Foster, 2003; Foster & Nascimento, 1994).

1.4 Relational color constancy

The traditional view has been particularly successful so far in the context of color constancy, where it is sometimes associated with the notion of 'inverse optics' (e.g. Maloney and Wandell 1986; Brainard et al. 1997; Brainard, 1998). Inverse optics is a methodology/research program which studies vision and visual phenomena under the assumption that the visual system generates our ordinary visual experience through the solution of inverse problems. In this framework, thus, color constancy amounts to the successful recovery, through an analysis of the retinal image, of whatever external feature corresponds to a surface's 'real' color, i.e. perceiver- and illumination-independent. In the next section, I will discuss an alternative account of color constancy that doesn't rely on the inverse optics paradigm, and in fact re-conceptualizes entirely what color constancy actually consists in. This alternative account comes out of a challenge to the

inverse optics modeling of color constancy, which I call "Foster's challenge". Even though the original challenge has a very narrow scope, namely to question whether the inverse optics approach to color constancy is satisfactorily supported by experimental results, I suggest that we draw a more general lesson from it, which opens up the possibility for a different understanding of *all* the constancies.

Even though different models of how the visual system recovers the perceiver-independent color of a surface, thereby displaying color constancy, have been proposed within the inverse optics tradition,[7] they all share the same idea: a retinal image – ambiguous with respect to which 'actual' colors (generally identified with Surface Spectral Reflectance properties, or SSRs[8]) distal surfaces have and how they are illuminated – must be first analyzed in order for the perceiver- and illumination-independent reflectance property to be isolated.

Even though it is undeniable that framing color constancy in terms of a solution to the inverse problem has so far been a successful working hypothesis, leading to the collection of a huge body of data on how the visual system *might* work when displaying color constancy, this doesn't automatically guarantee that the system *actually* works that way, or, more basically, that color constancy *actually* consists in the capacity to recover a specific SSR (or set of SSRs) while discounting other features such as illumination. Indeed, this very idea has been challenged.

Psychologist David Foster (2003) argued that most of the experiments done within the inverse optics framework would give the same results whether or not the visual system actually

---

[7] For example, Maloney and colleagues (1986; Maloney and Wandell 1986) propose a method to solve the inverse problem of color constancy which exploits the fact that there are more combinations of receptors on the retina than there are possible surface reflection profiles, while others (e.g. Golz and MacLeod 2002; MacLeod 2003; Brainard et al. 2006; Brainard and Freeman 1997) appeal to higher-order scene statistics and Bayesian models.
[8] E.g. Hilbert 1987; Byrne and Hilbert 2003; Tye 2000.

recovered perceiver- and context-independent colors of single specific surfaces by discounting the illuminant. Here is how 'Foster's challenge' goes. The most common experimental paradigm used by inverse optics researchers is the so-called 'asymmetric color matching'. In asymmetric color matching experiments, subjects are usually asked to adjust the perceived color of a patch so that 'it appears to be cut from the same piece of paper' as another patch. Both target patches are presented in an array with other, differently colored patches under two different illuminations (cfr. Figure 3).[9]

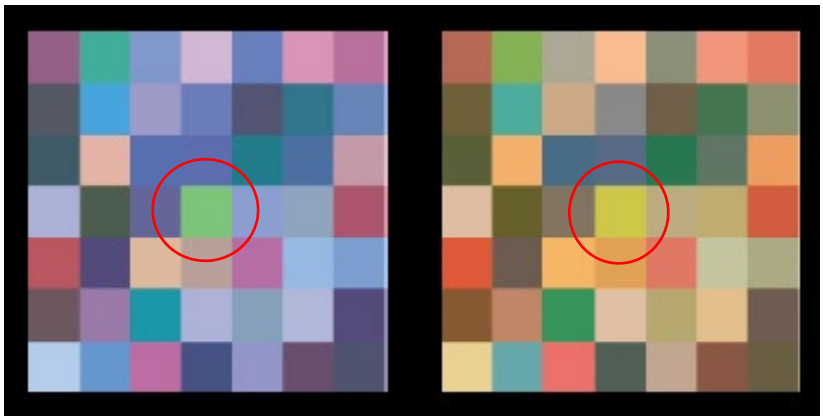

Figure 4. Asymmetric color matching.[10]

Traditionally, success in asymmetric matching experiments has been taken as evidence that subjects can successfully discount the effect of the illuminant and recover the 'actual' color of the patch. However, Foster pointed out that, given how the task is designed, subjects could succeed at

[9] When presenting the experiment's results, Foster warns that "it is unlikely that either will look like a match to the reader because they are not seen in the experimental conditions of controlled illumination." (2003, 440).

[10] In this experiment (image taken from Foster, 2003, p. 440), the two patterns of colored squares are identical, except for the fact that one on the left is illuminated by daylight (blue sky), while the one on the right is illuminated by the setting sun. Subjects were asked to adjust the color of the center square in the array on the right so that it 'looked to be cut from the same piece of paper' as the center square in the pattern on the left. This image represents a successful match.

matching the patches even without recovering their 'actual color' (or their SSR): it would be enough for them to rely on the way in which the patch in question *relates* to the other patches in the array with respect to color (e.g. this square is still more reddish than this other square, but more blue-ish of this other one, etc.). As Foster puts it: "to make a surface colour match, subjects need merely to judge how the colour of one surface relates to the colour of one or more other surfaces or, indeed, to the scene as a whole, first under one illuminant and then under another, or under changes in composition or configuration of the scene." (Foster 2003, 441). In short – assuming that success in asymmetric color matching experiments is a good indicator of successful display of color constancy – Foster's challenge suggests that the inverse optics approach is not the only one consistent with the data. Color constancy could consist in either one of two capacities: (a) 'traditional' color constancy (i.e. estimation and discounting of illumination + recovery of context-independent surface color), or (b) what's known as *Relational color constancy* (Foster & Nascimento, 1994; Foster, 2003; Amano et al. 2005, 2006; Zaidi 1998, 2002). The relational approach maintains that the visual system keeps track of whether the relations among the colors of multiple surfaces in the scene stay the same throughout changes in illumination.[11]

How to choose between the two theories? If we consider the question in the context in which Foster's challenge was originally formulated, the answer would be a totally empirical matter. It would in fact be equivalent to: Which one of the two models fits the data better? But Foster's challenge raises also a different, more philosophically significant question: which approach, if generalized beyond the color case, captures the nature of perceptual constancy better?

[11] In a more conciliatory spirit, Davies (2016) argues that color matching experiments support a pluralist view, where color constancy is sometimes traditional, and sometimes relational. I will not discuss Davies's view in this paper, but my main reason to resist pluralism is that, when we look at other constancies beyond color, the traditional view encounters difficulties (as the next section argues). This fact, together with a desire to provide a unified account of all the constancies across sense modalities, makes me inclined to put all my eggs in the relational basket.

It should be noted that, in fact, Foster himself claims that the relational view, rather than giving an alternative account of what color constancy consists in, merely shows that asymmetric color matching tasks are not the right kind of experimental paradigm to 'measure' color constancy. The fact that a relational interpretation fits the data is considered by Foster evidence that only something "weaker" than color constancy has been measured in humans. In this sense, Foster is still understanding and defining constancy according to the Traditional Constancy Thesis:

> In general, the challenge for the visual system is to estimate or eliminate the effect of the illuminant, so that surface reflectance can be deduced and, therefore, its true colour perceived. […] extracting information about surface reflectance remains a necessary condition for colour constancy. Without this information, colour constancy loses much of its rationale as a manifestation of a stable visual world." (Foster 2003, 439).

However, I think that Foster's challenge can be used to argue for something different, too. The simultaneous presence of these two models in color vision science is philosophically significant because each model can be seen as entailing a different view of *what* exactly perceptual systems *do* when they enable our ordinary experiences of invariance with respect to color. Because of this, the body of scientific research on color constancy offers us a concrete way to appreciate the different philosophical commitments of the traditional picture.

I will come back to the idea of relational color constancy in the second chapter, when I will use it to build an analogy with the case of auditory timbre constancy. Indeed, I believe that a general view inspired by the relational approach to color constancy is what I will be articulating and defending in this and the next chapters.

1.5 Conclusion

In this chapter, I showed that the way in which perceptual constancy is traditionally defined and understood is based on substantial theoretical assumption regarding how perceptual systems work and what their function is. Moreover, these assumptions seem to be challengeable once we articulate exactly what the capacity for constancy is supposed to do for us, namely providing a stable enough perceptual scene for us to judge accurately and act successfully. Many of the difficulties with the traditional view have been overlooked as a consequence of the popularity this view has enjoyed since constancy was first noticed and investigated. A new theory, thus, is way overdue. As I see it, an alternative approach to the traditional one should be more 'holistic', conceiving of proximal stimuli as involving facts about the perceiver as well as about what's 'out there', where the scene taken globally and the relations that exist within it are given a prominent role. Relational color constancy is a good starting point to understand what a non-traditional approach can look like and the kind of explanations it can provide. In the end, perhaps, the traditional approach will not be entirely replaced by the alternative. More likely, the two approaches will complement each other, and they might be needed *together* in order to understand such a multi-faceted capacity as perceptual constancy. As my work proceeds, it will become clearer both to myself and to my readers how the story will end.

At the start of this chapter, I pointed out that finding a 'neutral' definition of perceptual constancy, i.e. one that doesn't undertake any controversial theoretical committment, is a very hard, perhaps impossible, task. To define constancy *just is* to make certain assumptions regarding what the capacity for constancy consists in and what role it plays. The philosophers and scientists whom I take to be my opponents in this project define constancy as consisting in the recovery of specific perceiver-independent and local properties, isolated from the contextual features of the

perceptual scene as a whole, a task that is made very hard by the ambiguity and underdetermination of proximal stimuli. I have referred to this idea as the Traditional Constancy Thesis. In addition to it, I have formulated another thesis that contributes to the traditional picture: I called it *Constancy in Action*. The account of constancy represented by the two theses is much more controversial than it might appear at a first glance, and there are good reasons to challenge it.

I've argued against the traditional approach by giving possible counterexamples and suggesting different ways in which the empirical results cited in support of the traditional view could be understood. Moreover, I have pointed out that the capacity to judge and categorize objects in perception can come apart from the capacity to act successfully on the basis of perception. On the basis of this, I argued that it is at least premature to use evidence collected relying mainly on the former capacity (i.e. the majority of the evidence collected by the tradition) to support a view of perceptual constancy, as perceptual constancy seems to be involved in the latter capacity, too. As I will argue more explicitly in the next chapters, bracketing or rejecting the traditional approach opens up more interesting possibilities regarding what constancy might be. Much can be gained in terms of explanatory power, in terms of unifying instances of constancy in different modalities, and in terms of understanding the relationship between perceptual constancy and objectivity, on the one hand, and action, on the other.

2.0 Non-visual constancy

2.1 Introduction and background

I have presented the traditional view of constancy in chapter 1, and my discussion of constancy in sense modalities different from vision in this chapter takes for granted the reconstruction of the traditional view I gave back then. As a reminder, here are the two theses which I think summarize the traditional approach to perceptual constancy:

Traditional Constancy Thesis: Perceptual constancy is the capacity of perceptual systems to recover the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli.

Constancy in Action: Recovering the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli is necessary for successful perceptually guided motor action.

Moreover, in chapter 1 I have pointed out that the way in which the tradition generally takes specific perceiver-independent properties to be recovered and attributed to objects is via the solution of *inverse problems*. In what follows, I will assume that my opponent, i.e. the tradition, is characterized in terms of the two theses above, plus the further commitment to the 'inverse-problem strategy'.

The current chapter will be dedicated to motivate and defend the following claim: constancy is a capacity that has both visual and non-visual 'versions', because the same minimal results are obtained and the same role is played in all cases. This claim is rejected by the tradition on the basis of the definition they rely on (which we essentially owe to Burge). If my defense of the claim is convincing enough, then, I will have established the need to look for a new approach to perceptual constancy. The capacity for constancy, according to the alternative, does not involve

the extraction/representation of specific perceiver-independent properties. Notice that I am *not* claiming that the visual or other sensory systems lack altogether the capacity to form mental representations of the actual properties of objects. This would be false. What I am claiming, instead, is that such a capacity is *different* from the capacity for constancy, and that we should not define or understand the latter in terms of the former, as the tradition does.

In section 2, I will present two cases in which, I argue, constancy is displayed in the context of *auditory* experience. With the help of some empirical evidence, I will argue for three claims. First, these auditory cases have a lot in common with visual displays of constancy. Second, the traditional approach cannot straightforwardly include these auditory cases as displays of constancy, because they don't seem to conform to their definition (i.e. no inverse problem seems to be solved, and no specific perceiver-independent property seems to be recovered). Third, in light of the two previous claims, we have one more good reason to be skeptical about the traditional view and to start looking for an alternative. Looking at non-visual cases of constancy will allow me to start sketching my preferred approach to constancy, even though a more thorough discussion of my positive proposal will have to be postponed to a later chapter. In section 2.3 I will provide further considerations in favor of the conclusions I drew in section 2.2.

2.2 Constancy in non-visual modalities

In this section, I discuss a couple of possibilities with respect to the existence of a version of the capacity for constancy in a non-visual modality, namely audition. The cases I present aim at supporting the following suggestion: there is no good reason to limit the talk of constancy to vision and to whatever the tradition has said constancy consists in. Scientists studying non-visual

perception acknowledge that non-visual systems have the capacity to track invariant features in the environment, that is, such systems have a capacity that relates to the capacity for constancy just like these invariant features relate to the perceiver-independent local properties mentioned in the Traditional Constancy Thesis.

Most scientists claim that invariants are non-visually tracked/represented, yet they do not tend to see this capacity for tracking invariants as analogous to visual constancy. I believe that perceptual modalities other than vision have been considered incapable of constancy because of the specific way in which constancy is defined and approached within the tradition. Most scientists studying audition, smell, taste, etc. still agree that vision is 'special' in the way in which it seems to precisely 'pick out' the intrinsic properties of objects in the world. It is surely tempting, therefore, to just define constancy, which after all is a capacity for stability of the perceptual scene, as precisely the capacity to pick out such properties, as they are stable in the sense of independent of any change in either the context in which they are experienced or the perceiver.

Despite being an easy and tempting way to understand constancy, however, it is wrong, and cases of constancy being displayed in non-visual modalities are good evidence for this wrongness. All versions of perceptual constancy, visual and non-visual, ultimately play the same role for the creature: they enable the creature to experience a perceptually *stable* world in which she can successfully act. In this respect, I will argue that a visual version such as color constancy and an auditory one such as timbre constancy (which underlies timbre identification) have quite a bit in common. But let's present the cases in a little more detail first.

2.2.1 Auditory constancy, part I: loudness

Zahorik and Wightman (2001) tested to what extent listeners can experience a sound as maintaining the same loudness when the distance between them and the sound source changes, that is, tested whether there is such a thing as *loudness constancy*. It turns out that, on average, loudness constancy is pretty robust in this situation. The goal of the experiment is to provide an account of loudness constancy which does not obviously depend on the solution of a 'standard' inverse problem in which a specific sound intensity (generally associated with the amplitude of the sound wave) is recovered and attributed to the sound, in a way that would be analogous to the traditional account of visual size constancy with varying distance.[12]

In the case of visual size constancy, the traditional account states that an estimation of distance must be extracted from the ambiguous retinal stimulus together with an estimation of actual size. Size constancy with varying distance, therefore, is displayed, according to the traditional approach, in virtue of the attribution of the estimated size-property to the object, 'discounting' the previously recovered estimation of distance together with the rest of 'perceiver-involving' variables. However, the analogy between size and loudness doesn't quite work. First of all, according to the authors, the data show that subjects are quite bad at estimating distance from the source. At the same time, however, their capacity for loudness constancy remains pretty good (see Figures 5 & 6).

[12] Cfr. Sperandio and Chouinard (2015) for a review of current accounts of size constancy. The hypothesis is that both retinal image size and distance from the observer need to be processed in order to achieve size constancy.

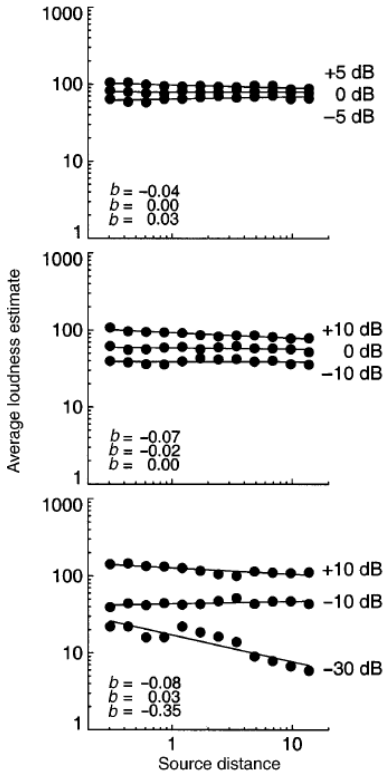

Figure 5. 'Objective' loudness judgments at different distances.

The graph in Figure 5 (from Zahorik & Wightman, 2001, 79) illustrates how, when asked to estimate 'objective' loudness of a sound at the source, listeners are good at compensating for changes in the distance between them and the sound source. This is taken by the authors to be preliminary evidence that the auditory system capable of loudness constancy with varying distance. By themselves, these data don't support any particular theory of what exactly loudness constancy is and how it works. It is just assumed that these data 'prove' the presence of loudness constancy, whatever it is.

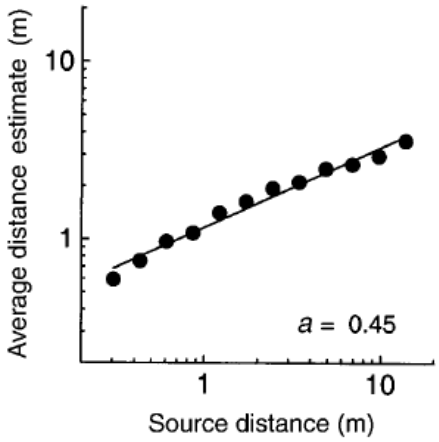

Figure 6. Estimated vs. Actual distance of the source.

In figure 6 above, we can see how listeners tend to judge a sound source to be much closer than it actually is. This suggests that 'objective' distance estimation is not necessary for loudness constancy. In turn, this claim goes against the traditional interpretation, according to which loudness constancy depends on the auditory system explicitly estimating and then discounting information about source distance in order to recover 'objective' loudness at the source from the proximal stimulus.

In light of these results, the authors argue that correct distance is unlikely computed, whether consciously or not, and that, on the other hand, estimating distance itself cannot be among the *necessary* conditions for the display of loudness constancy. This, in turn, explicitly contradicts the standard traditional approach, which takes every instance of constancy involving changing distance to necessarily include distance as a 'variable' in the 'equation' representing the inverse problem.

However, here is an important thing to keep in mind while reading about this experiment. Despite denying that direct distance estimation is necessary for loudness constancy to be achieved, the authors still admit that, if what they are studying is indeed loudness *constancy,* then it must fit

the traditional definition and experimental methodology. Indeed, what they take to 'prove' that loudness constancy is rather robust across changes in the location of the sound source is the fact that subjects seem to be accurate at judging the specific intensity of the sound at the source. Accurate estimation of intensity at the source is, for them, the sign that loudness constancy is being displayed, because in the traditional framework intensity at the source would precisely be the perceiver-independent property that needs to be recovered in this case. However, the fact that the experimental set-up is still traditional doesn't affect my broader argument, because I use this experiment only to suggest that even if one considers the capacity for constancy to be a matter of recovering specific perceiver-independent properties (in this case, intensity at the source), in the case of loudness constancy the standard inverse-problem strategy doesn't seem to be employed. And without an inverse problem being correctly formulated, the traditional approach cannot explain where the recovered property 'comes from'. [13]

The account of loudness constancy proposed by Zahorik and Wightman, despite still being 'traditional' in the sense that it relies on the accurate attribution of a perceiver-independent property like intensity at the sound-source, nonetheless has a relational 'flavor' to it. Indeed, if the auditory system were sensitive to the *relations* between the physical properties of the sound waves and the physical properties of the environment in which they travel, we could explain how intensity at the source can be accurately recovered *without* appealing to a strategy that requires distance estimation first. Zahorik and Wightman's proposal goes precisely in this direction. They suggest

[13] The authors still admit that an explanation in terms of inverse problem factorizing the proximal stimulus into intensity at the source and distance and then 'discounting' the latter is still possible. However, they argue that such an explanation would have to make some commitments that might be impossible to verify with our current psychophysical methods. Ultimately, therefore, an explanation that doesn't require such a 'leap of faith' would still be preferable, and this is what they aim to offer.

that the auditory system might be tracking a *relational* property of the auditory scene (sound + contextual environmental features) called "reverberant sound energy".[14]

Reverberation is the persistence of a sound in a closed environment (e.g. a room, but also a cave, or a thick forest) due to the soundwaves produced encountering material surfaces that partially absorb, partially reflect them. Compared to echo, i.e. another acoustic effect due to soundwaves 'bouncing' into surfaces, reverberation occurs when waves are reflected from multiple surfaces in multiple directions at the same time, and much more quickly than in the echo case, due to shorter distance between the sound source and the reflecting surfaces. Reverberation depends on frequency, both of the original soundwave and of its reflection, and is a powerful acoustic cue for spatial features of the environment in which the sound has been produced, as well as, or so Zahorik and Wightman argue, for estimation of intensity at the source. Indeed, the intensity of the reverberation produced in the same environment by a sound whose source changes its distance from the listener is directly proportional to the intensity of the sound at the source, i.e. their ratio remains invariant while distance between source and listener varies. This invariant ratio, then, the authors suggest, would enable listeners to estimate source intensity – when asked to do so in the experimental context – without necessarily estimating distance first. In this picture, the proximal stimulus cannot be taken to be constituted exclusively by sound intensity at the ear, but it must contain reverberation cues, too, and thus cues regarding the auditory scene as a whole. The proposal, thus, is that loudness constancy might depend on the auditory system's sensitivity to invariant reverberation, which is a relational property of the auditory scene as a whole (it is, indeed,

[14] Predictions of distance estimation based on the hypothesis that listeners are directly tracking reverberant sound energy in their environment are, indeed, consistent with the graph in figure 1.b.

defined as the constant ratio between sound intensity at the source and decay rate after the waves are reflected by surfaces).

Despite the fact that Zahorik and Wightman still adopt a large part of the traditional framework (i.e. they equate loudness constancy with the capacity to estimate perceiver-independent intensities at distal locations), the tendency to move away from the inverse-problem strategy in non-visual perception science is nonetheless good news. Recall that, as I mentioned in chapter 1, something similar is happening in vision science, with the relational color constancy proposal (e.g. Foster and Nascimento 1994; Zaidi 1998, 2001; Foster 2003; Amano et al. 2005; Nascimento et al. 2005). In the next sub-section, I will examine another version of auditory constancy for which an explanation in terms of inverse problem, and this time also in terms of recovery and attribution of a specific perceiver-independent property, seems quite difficult.

2.2.2 Auditory constancy, part II: timbre

For audition scientists, timbre has moved from being considered a quite straightforward acoustic property, entirely reducible to the shape of the corresponding soundwave (e.g. Helmholtz 1885/1954) to being an elusive 'chimera'. Indeed, timbre is now universally thought to be a quality that, if definable at all, it can only be defined along multiple dimensions, standing in very complex relations with each other (Elliot et al. 2013; Siedenburg and McAdams 2017) and, consequently, hard to model as a property of the sound-event taken in isolation from its context. From the point of view of experience, moreover, it turned out to be very hard to provide a definition of timbre that goes beyond its negative identification as 'not pitch and not loudness' (Bregman 1990).

It is now commonly thought that, when listeners identify sounds and recognize them on the basis of their timbres, they do so reliably enough for scientists to use their responses to model

potential candidates to be the physical correlate for timbre, even though such a correlate is way more complex than the 'mere' waveform. Despite the difficulties deriving from the multi-dimensionality of timbre's physical correlate, the attempts to isolate it are motivated by the fact that we seem to show a quite impressive degree of stability of experience across changes in listening context which, in turn, affect the physical properties of the waves reaching our ears. Philosophers are increasingly taking interest in the 'puzzle' of timbre identification precisely for this reason. And talking about timbre *constancy* seems to me not just useful in this context, but probably *necessary* in order to fully understand what's going on. Here is Isaac (2017):

> Anecdotally, we perceive timbre to remain constant across changes in the source that affect the acoustic properties of the signal, i.e. its spectral composition. The soundwaves from a violin played in the same room as the listener, from one played down the hall, and from one played outside the window will be quite different in spectral composition, yet we perceive them as ''the same,'' or at least very similar, in timbre. (513-514).

Both the tradition and its opponents, I think, would read in the passage above a hint that the *explanandum* is the capacity for perceptual constancy with respect to timbre: the stimulus at the ear changes drastically over time, due to the soundwaves' interactions with different surfaces and media as they travel, but we have the capacity to experience the sound as stably maintaining its timbre nonetheless. The tradition, however, might disagree, because, as I have mentioned, it is hard to identify precisely which specific property attributable to something distal corresponds to our experience of timbre. Not only is it multi-dimensional and therefore complex to recover in all its components, but it might also be the case that information about timbre is carried by different combinations of variables at different times and in different listening contexts.

The shared hypothesis is that the auditory system, when it allows us to identify and recognize sounds on the basis of their timbre, does so by being sensitive to a complex pattern of regularities that show up over time and are 'spread' throughout the auditory scene (or

‘soundscape’). In this ‘constellation’ of variables that is hypothesized to be the ‘substrate’ of timbre, the physical properties of soundwaves certainly play a role, but aren’t the only contributors, and it is hard to characterize the role both quantitatively and qualitatively once and for all. Roden (2010, quoted in Isaac 2017) puts the idea in the following terms:

> Timbral discrimination […] does not plausibly ‘track’ a single type of physical feature…but relatively idiomatic patterns of relations between such features. This is consistent with a qualified interpretation of timbral kinds as consisting of recurrent constellations of features of sound generation processes, but it need not entail an essential limit on what kinds of relationships between more basic physical features can be picked out through identification of timbres. (145)

In other words, our own experience tells us that have the capacity for timbre constancy, while audition scientists tell us that we cannot have the capacity to recover and track a fixed distal invariant (or set of invariants), because such a fixed distal invariant correlating with timbre might not actually exist. Timbre constancy, therefore, is more likely to depend on the auditory system’s sensitivity to regularities emerging over time and largely dependent on the relations and interactions occurring within the soundscape. Moreover, we should not exclude the possibility that these regularities necessarily manifest themselves in virtue of certain features of the auditory system itself, such as its degrees of adaptation, its over-sensitivity to change, etc. This would mean, in turn, that timbre constancy might be intrinsically perceiver-dependent in a way that is at odds with the traditional understanding of how constancy works. But let’s proceed one step at a time. Let’s bracket the issue of perceiver-dependence for now and focus on clarifying why I think that the tradition is in trouble when trying to account for timbre constancy. The Traditional Constancy Thesis applied to timbre would sound more or less like this:

Traditional Timbre Constancy Thesis: Timbre constancy is the capacity of auditory systems to recover an absolute, perceiver-independent property (or cluster of properties) of a sound-event, by discounting contextual and perceiver-dependent elements in proximal stimuli.

But we saw that, in the case of timbre, audition scientists do not believe that any specific property (or fixed cluster of properties) of the sound-event, the recovery of which would allow the auditory system to display constancy in a way compatible with the traditional definition, can be isolated. Rather, a plausible alternative seems to be that listeners recognize sounds on the basis of timbre by being sensitive to non-reducible regularities that emerge from the global features of the soundscape and the relations among elements within it. This non-reducibility of timbre to perceiver-independent properties of sound-events is incompatible with the traditional definition of timbre constancy.

As a matter of fact, we *do* seem to have at least the skeleton of an alternative approach: we could take the relational approach to color constancy and 'extend' its main insight beyond color. Interestingly, I am not the only one who has suggested to explore this possibility. Stilp et al. (2010), for instance, argue that the capacity that enable sound identification on the basis of timbre has quite a lot in common with color constancy understood according to the relational model:

> Analogous to vision, spectral composition of sound entering the ear is colored by the listening environment. Energy at some frequencies is emphasized by acoustically reflective properties of surfaces, whereas energy at other frequencies is attenuated by acoustic absorbent materials. In this way, listening context spectrally shapes the acoustic signal. […] effects of listening context are, in fact, closely analogous to visual color constancy. (471)

They tested the hypothesis in a series of experiments, though here I will only discuss one of them. They had subjects listen to a series of six synthetized sounds that gradually changed from being very similar to a French horn in timbre, to being very similar to a saxophone. Before playing the target sounds and asking the subjects to decide whether, based on timbre, the sound was being produced by a French horn or by a saxophone, the experimenters imposed what they take to be the

auditory analog of a 'light filter' (e.g. a transparent red film put in front of a light source[15]) on the listening context. Such 'auditory filters' consisted in a series of non-speech sounds which shared some spectral characteristics with sounds emitted by a French horn, by a saxophone, or several combinations of both. They predicted that listeners would report hearing a saxophone more often when the preceding sound shared some spectral characteristics with the sounds normally emitted by a French horn, and, vice versa, they would report hearing a French horn more often when the preceding sounds were more similar to the sounds normally emitted by a saxophone.

The general aim of the experiment was to show that the auditory system keeps track of salient changes in proximal stimulation and simply 'assumes' that regularities that are spread over the whole auditory scene (like a light filter affecting the illumination of all the objects in the room, indeed) are 'irrelevant' with respect to the task at hand – i.e. isolating and tracking a unified auditory object on the basis of timbre. The results (see figure below) seem indeed consistent with the hypothesis that timbre constancy works somehow similarly to relational color constancy.

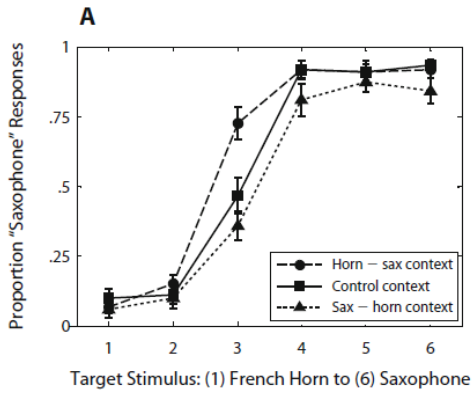

Figure 7. From Stilp et al., 2010, 477.

[15] "The spectrum of illumination can be considered to be a filter that is reliably imposed on the full context of viewing, and perception of constant color is maintained by relative differences between the spectral composition of the object being viewed versus reliable spectral characteristics of the viewing context" (Stilp et al. 2010, 478).

Even though these results don't in principle rule out the idea that timbre constancy could work in a more traditional way, some further considerations about the nature of the auditory system might be more consistent with a relational approach. In particular, the way in which we know the auditory system processes information seems to suggest that some sort of perceiver-dependence should be taken into account when thinking about constancy in audition. Minimally, Stilp and colleagues' study shows that auditory perception can be explained without much difficulty within an approach to perception in general where the stability of the perceptual scene depends on relations and global features of the whole scene, to which the system is directly sensitive (analogous to the sensitivity of the visual system for contrast). Something along the same lines has been proposed regarding color constancy, on the basis of the idea that color, just like timbre, might be much more complex and multi-dimensional property than the tradition initially thought (e.g. Cicerone et al. 1995, Wollschläger et al. 2001, Brown 2003, Hoffman 2003).

Kiefte and Kluender (2008) argue that the auditory system is directly sensitive to relations and patterns emerging within a temporally extended proximal stimulus, and shows the capacity to adapt quite quickly to regular and 'predictable' features affecting the soundscape as a whole (in the same way in which illumination affects the whole visual scene in the case of color perception). By quickly adapting to global features and treating them as 'background', "perceptual systems respond predominantly to change. They do not record absolute levels whether loudness, pitch, brightness, or color […]" (374). Focusing in particular on the auditory system, the experiment described above is consistent with the idea that the system quickly adapted to the predictable global features of the soundscape (just like the visual system does with light, at least in the relational color constancy model) and had an enhanced response to anomalies in the relations within the auditory scene as a whole. If no anomalies arose, and thus the internal relations were maintained

unvaried over contextual changes over the whole scene (i.e. different 'filters') the result was a display of constancy.

To sum up, I have suggested that timbre constancy plays the same role as other, more 'standard' versions of perceptual constancy (e.g. shape or color) in providing stability of the perceptual scene so that the perceiving creature can act (including issuing judgments to identify or recognize) successfully. Second, if talking about timbre constancy is plausible enough, then it seems that the traditional definition is inadequate, because it cannot be straightforwardly applied to this case. Indeed, a definition of constancy that relies on the recovery of specific perceiver-independent, local and distal properties can account for something like timbre, which is more likely a dynamic and non-fixed combination of different features of the entire soundscape, perceiver and her auditory system included. Third, I have suggested to understand timbre constancy in terms of adaptation to global features acting as 'filters' and sensitivity to relations within the soundscape that may, on the contrary, be salient for action and identification of sound sources. We saw in chapter 1 that, according to some, something similar happens in vision, when sensitivity to color relations within a visual scene, against a background of constant illumination, are taken to be at the origin of color constancy.

Before moving on to the next section, I think I am now in a position to list some preliminary features of the approach to perceptual constancy that I favor. This approach is different from the traditional one in a similar, though not fully analogous, way in which relational color constancy is supposed to be an alternative to the 'inverse optics' paradigm. The approach to perceptual constancy emerging from my discussion of non-visual cases has, for now, at least the following characteristics:

1) It aims at understanding the capacity for perceptual constancy in terms of its *practical role* in the context of action coordination and the perceptual stability needed for it.
2) It rejects the idea that the capacity for constancy is identical to the capacity to recover specific perceiver-independent and local properties attributable to distal objects or sets of objects.
3) It rejects the inverse-problem strategy.
4) It understands proximal stimuli in a more 'inclusive' way, in which global features of the scene (and the *relations* among them, including the perceiver) are *constitutive* of it. The proximal stimulus does not need to be decomposed, and the 'real' invariant (i.e. perceiver-independent property) does not need to be picked out.

2.3 Tying up a few loose ends

In this chapter, I have tried to emphasize that there are visual and non-visual versions of the capacity for perceptual constancy. These versions have a lot in common. In particular, I have argued that, once we take a 'pragmatic stance' and we look at what perceptual constancy *does* for a creature, there is no actual reason to treat visual and non-visual version of the same capacity differently. The only reason why non-visual constancy hasn't been fully acknowledged so far is, I think, the fact that the traditional definition of constancy has been the orthodoxy so far, and such a definition does not easily apply to perceptual modalities other than vision (and, probably, even to vision itself). I also started sketching what I think is a good, though not yet completely satisfactory, way to unify our account of visual and non-visual constancy. The alternative framework that I have in mind, indeed, takes as a first starting point the relational approach to

color constancy proposed by Foster, Zaidi, and others, and extends the insight that lies at the core of this approach to other versions of constancy. Here are some further examples.

In haptic touch, we seem to be capable of roughness constancy when touching a surface throughout changes of the speed at which our tactile receptors move along the surface (e.g. Yoshioka et al. 2011; Boundy-Singer et al. 2017). Although it is controversial whether the proprioceptive feedback deriving from actively moving the receptors along the surface is needed for this constancy effect, it seems a stretch, given the nature of touch as a sense modality, to say that the tactile system is solving an inverse problem and is computing the actual roughness of the surface out an underdetermined proximal stimulus. A more relation-based approach, on the other hand, can account for roughness constancy in terms of the capacity of skin's receptors to adapt to and 'level out' changes irrelevant for the current behavioral goal.

One more potential example comes from olfaction, which has been taken, especially by Burge, as the paradigmatic sense modality which is incapable of constancy. I don't see any reason not to consider displays of constancy the well-documented cases in which 'olfactory objects' are tracked quite robustly through time as well as through changes in odor intensity either of all components or even just of a single one of them (e.g. Barnes et al. 2008, Gottfried 2010). He only thing getting in the way is the traditional approach. This olfactory capacity is known as 'concentration invariance' (Uchida and Mainen 2008; Bolding and Franks 2018), and it has sometimes been suggested, indeed, that concentration invariance is in fact olfactory constancy (Asahina et al. 2009, Gottfried 2010). The tradition could not accept this, because both psychological and neurophysiological studies have shown that tracking of complex olfactory objects does not require the recovery of any specific perceiver-independent odor property. Rather, olfactory systems seem to be naturally disposed to keep track of certain internal relations among

the odor components that together give the olfactory object its internal coherence. Olfactory objects, in other words, are tracked through intensity changes by focusing on the unity of the odor as a whole and on the relations among the various odor components remaining fixed (similarly to what happens in relational color constancy with changing illumination).[16]

Sensory modalities like olfaction, kinesthesis, proprioception, or haptic touch, have been considered 'non-perceptual' by Burge and the tradition because incapable of going beyond their 'narcissistic' (Akins 1996) function of directing sensory input inside the body rather than towards the external world. Nevertheless, these modalities make it possible for the perceiver to act successfully on the basis of their input, as a consequence of their 'stabilizing' function. I am aware that more needs to be said in favor of my invitation to depart from the traditional approach, because such an approach has been taken for granted for a long time and it permeates most of perceptual science, not only of vision. In the next subsections, I will discuss a few issues that might indirectly make my proposal to 'think outside the traditional box' a little more attractive.

2.3.1 How to choose a theory: quality vs. quantity

Proponents of the traditional approach of constancy might have grounds to claim that, if their definition cannot include cases like those discussed in section 1 here, then those cases should not count as displays of constancy, period. After all, their definition has served as the basis for a huge amount of successful experiments, fairly accurate predictions, and, most of all, a very high level of *understanding* of how perceptual systems, and especially the visual system, might process

[16] For further critiques of the traditional framework based on olfactory perception, see Barwich (2019, 2020) and Smith (2015, 2017).

information. I stress the word 'understanding' because, in evaluating whether the alternative, relational approach is better equipped than the traditional one to explain perceptual constancy and its role with respect to action, we must take into account one issue that might arise when scientific theories and the systems they aim at modeling/explaining get more and more complex. Often, the level of understanding, or intelligibility, provided by a psychological theory or framework (mostly due to its abstract and simplified nature) tends to be inversely proportional to its capacity to model phenomena in a more fine-grained way, thus issuing more accurate *predictions* relative to those phenomena. A highly intelligible framework, that is, one that scientists can almost fully comprehend in the way it works and issues its predictions, generally issues predictions that aren't extremely precise, but, on the other hand, it provides explanations in wider variety of contexts and circumstances – it is, in other words, more *flexible*. Flexibility, however, comes at the price of abstracting away from many details in order to obtain a broader range of application, and this, in turn, means giving up, at least to an extent, ecological accuracy.[17] When comparing the traditional and the relational approaches to perceptual constancy, therefore, we have to make sure we take into account the necessity of a 'trade-off'[18] between different scientific or philosophical 'virtues' and explanatory goals that might stand in the way of one another, in the same way as prediction and understanding, or ecological accuracy and flexibility, do in a number of other cases.

---

[17] How we should characterize the relation between the degree of simplification and abstraction of a scientific model and its capacity to be intelligible to scientists and, in turn, to enhance understanding of how the world works, is a hotly debated question in philosophy of science (e.g. Khalifa 2017, Chirimuuta, unpublished manuscript). The crucial point of disagreement turns out to be the nature of understanding, and in particular whether we should consider understanding factive (i.e. we don't understand a claim x unless that claim is true) or non-factive (i.e. understanding is independent of truth). If we take understanding to be factive, this is in clear tension with the idea that the more simplified and distorted by abstraction and removal of 'unnecessary details' a model, the more intelligible and understanding-promoting. After all, a simplified and distorted model is almost certainly false of the system it's modelling. On the other hand, however, if understanding is non-factive, then simplification and abstraction in a model can legitimately be understanding-promoting features: whether a model is 'saying' something true of the system it's modelling becomes irrelevant, as long as the model is still saying something about the system that helps scientists understand at least some fundamental features of it.

[18] See previous footnote.

What are the 'virtues' and advantages that characterize each approach in comparison with the other? Recall that the traditional approach starts from an analysis of the perceptual scene and its 'decomposition' into simpler features. The features relevant for constancy (i.e. the specific perceiver-independent properties of objects or locally identified combinations of objects) are then kept, while the rest is discarded. This approach has been quite successful for centuries in, on the one hand, issuing 'good enough' predictions and, on the other hand, enhancing the understanding of the capacity for perceptual constancy. Especially with respect to visual cases like color or shape constancy, the tradition has created a number of computational models and collected data through their application. These models created within the tradition are generally constructivist, or inferentialist, in spirit. A good example is constituted by the models of color constancy created within the 'inverse optics' paradigm (cfr. ch.1, sec.2). Put crudely, the traditional approach to constancy chooses a trade-off that favors quantity over quality in terms of producing large bodies of data in virtue of the flexibility of the model to 'fit' different tasks (and biologically different visual systems). On the other hand, an approach that focuses on relations and takes the scene in its entirety as the level at which perceptual constancy displays itself and can be fully understood, tends to be more 'quality-oriented'. This means that the approach I favor might increase the complexity of the models it creates and, probably, reduce their flexibility as a result.

In sum, a theory that is quantity-oriented, flexible, simple, and abstract, is also usually more intelligible. Intelligibility, in turn, promotes the growth of the body of experimental research relying on this theory. Hence the success and popularity of the traditional approach, both to constancy and to perception in general. On the other hand, the approach that I started sketching, which has at its core relations instead of properties, and is more 'holistic' and less 'analytic', would have the advantage of being more ecologically accurate with respect to what we know about the

nature of perceptual systems in different sense modalities, as well as across different species. Even though the models created within the relational approach will have to incorporate different details depending on certain ecological and biological constraints, such as, e.g., the mechanisms characterizing auditory as opposed to visual perception in humans, such an approach has the potential to unify the notion of perceptual constancy in a theoretically quite powerful way. In particular, the notion of *stability for action* combined with a relational and holistic framework elucidates what instances of constancy across modalities and across species have in common.

Before concluding, I will now turn to a possible way in which the tradition could come back and unify constancies within their framework, too. In particular, they invoke Bayesian models of perception. I will argue that a 'Bayesianized' version of the traditional approach, at least in the context of perceptual constancy, is still inferior to the relational one I defend.

2.3.2 What about Bayesianism?

One might point out that Bayesian models of perceptual processing (e.g. Clark 2012, Hohwy 2013, Rescorla 2015, Gładziejewski 2016) are attempts to unify perception in different modalities (and perhaps even the whole 'mind'[19]) without abandoning much of the traditional conception. Seeing the brain as performing Bayesian inferences to determine and integrate the stimuli and, consequently, the higher-level perceptual states, indeed, seems to be a way to introduce a 'pragmatic' element, through the way in which priors are formed, for instance. Whether this approach can actually be an improvement of the traditional approach with respect to perceptual constancy, however, remains to be seen. What would it mean to 'Bayesianize' the traditional

[19] See, e.g., Danks (2014) and Clark (2015).

approach to constancy? Roughly, it would mean that the way in which the recovery of specific perceiver-independent properties from the underdetermined proximal stimuli follows a procedure similar to a Bayesian inference.[20] For example, the recovery of a certain specific color property within an asymmetric color matching task, in a Bayesian framework, would require the system to take into account relevant sources of 'evidence' and determine the color property to be attributed to the surface as the conclusion of an inference where the result is the most probable given the 'evidence' (including past experience, which can affect the *likelihood* variable in the relevant equation).

Does making the recovery procedure Bayesian instead of based on some other kind of strategy (e.g. more classic rule-based computations) help the traditional approach? Does it make the approach more plausible? I don't think so, and here is why. First of all, even though the Bayesian framework is often invoked as a way to revise a traditional conception of perception (and cognition) by making it less 'isolated' from environmental factors, past experience, etc., is still a framework that assumes percepts to be the result of *inferences.* This idea is itself not different from the non-Bayesian kinds of traditional accounts: the inferential process might include more or less variables, and might weigh those variables and relate them in different ways, but it is always an inference we are talking about. Bayesianism, in other words, is a methodological tool that doesn't touch on the basic assumptions of the traditional approach, which are, however, the ones I am contesting. The disagreement among theories of perception, and in particular among approaches to perceptual constancy, that I am interested in is on another level: namely, the level of *what* is being computed by the perceptual system: perceiver-independent, localized properties or relations

[20] Cfr. Purves and Lotto (2003).

within entire scenes? That's the issue with respect to which I am contrasting the traditional approach with a relational one, and the popular Bayesian framework has very little to say about it.

Moreover, notice that Bayesian models are importantly *idealized,* because they aim at capturing *optimal* decision-making processes. Perceptual systems (as any other biological system probably) are simply not the kind of things that is even remotely optimal. Bayesian modelers are aware of this, and indeed, whenever a Bayesian account of a perceptual process is proposed, assumptions that aren't biologically correct need to be made (Knill and Pouget 2004, Chirimuuta 2018). If the question is whether 'Bayesianizing' the traditional approach to constancy, the way in which, for example, Brainard (1998, Brainard et al. 1997, Brainard et al. 2006) does in the context of color, takes away the need for a relational and 'holistic' alternative, I think the answer is 'No'. Bayesian models are idealizations of perceptual processes, rooted in biologically implausible assumptions, which, moreover, do not solve the problem that I have pointed out with the very idea of recovering certain perceiver-independent properties and to do so accurately. Changing the *procedure* (making the inference Bayesian instead of non-Bayesian) with which perceiver-independent properties are recovered still assumes that recovery of perceiver-independent properties is what the system does or aims to do. If, according to advocates of the tradition, the goal of a 'Bayesianization' of their approach to constancy, which is already quite predictively powerful, is to make it more ecologically accurate, then the fact that Bayesianism relies on biologically false assumptions immediately rules that out. If, on the other hand, the goal is to make the traditional approach fit non-visual cases better, which is, as I have been arguing, an advantage of the relational approach, then 'Bayesianizing' doesn't help, because what is needed in order to accommodate non-visual cases is to abandon the idea that the system is recovering (via some kind of pseudo-inferential procedure) a specific perceiver-independent property of the world, and a

'Bayesianized' version of the relational approach is still traditional (i.e. pseudo-inferential) in all the relevant senses.

2.3.3 The 'scene-parsing' hypothesis: tradition 2.0?

Another somewhat recent proposal worth mentioning is the one by Mohan Matthen (2010). Matthen clearly acknowledges the need to overcome a strictly atomistic approach to perceptual constancy: attributing certain local perceiver-independent properties to specific objects cannot be the whole story. In particular, he argues in favor of what he calls the "Scene-parsing perspective":

> When sensory receptors are excited by signals determined by several environmental objects varying independently in their properties, the function of the sensory system is to extract from the state of the sensory receptors an assignment of properties to all such environmental objects as are of interest to the organism. (Generally, it does so by exploiting existing environmental correlations found in *scenewide* statistics.) (2010, 238; emphasis in the original)

Matthen argues that, when we try to explain constancy at a phenomenal level, we are initially confronted with an apparent puzzle. Take the experience of color constancy of a wall different portions of which are illuminated differently: clearly, there is a sense in which the wall looks of a uniform color, and there is a sense in which the different portions of the wall look different in a way that definitely involves color. Thus, one might ask: does the wall look uniform in color or does it look as having different colors in different areas? Answering 'both' would obviously be inconsistent, but Matthen points out that such an answer nonetheless captures something true about the phenomenology. It is apparent in our visual experience *both* that the wall maintains its color throughout *and* that it looks different in a color-related way at different locations. The logical inconsistency, turns out, is something that can be easily eliminated by

noticing that it is not mandatory, in order to have a constancy experience of an object, that all the properties playing a role in the constancy experience are attributed *to the object*:

> My desk lamp has a yellow shade, and it is casting a yellow patch on the pastel-colored wall beyond it. This makes it difficult to see the color of the wall. But it is, at the same time, visually perfectly evident to me that the color of the light cast by the shaded lamp is not the color of the wall. […] Of course, the fact that the spot of yellow light obscures the color of the wall can lead to errors of perception regarding the wall. But this does not mean that it modifies the color of the wall, or makes it look as if it has a different color. (249)

Sometimes, for instance, color constancy with respect to a specific object can involve the attribution of properties to other items in the scene, such as the light. In this way, color constancy becomes, in fact, a 'problem' solvable only by looking at the scene as a whole. Matthen's proposal is, once again, a step in the right direction: he recognizes that some kind of holism about perceptual scenes and perceptual experiences is necessary in order to make sense of constancy. However, he is still very much committed to traditional conception, with at the center perceiver-independent properties and their attribution to specific parts of the scene. In other words, in Matthen's account the *relational* component with the perceiver is still missing. As I have tried to argue throughout the chapter, a better understanding of constancy can be provided by setting aside both an overly atomistic conception of perceptual processing *and* the idea that constancy relies on the attribution of specific properties to specific objects. This seems particularly necessary if our goal is to unify visual and non-visual versions of constancy under the same general explanatory framework.

2.4 Very brief conclusion

Providing an explanatory framework that can unify a set of capacities which seem to share the same functional role in the life of a creature is, I think, an intrinsically valuable philosophical

task. In the particular case of perceptual constancy, I have argued that, once we take the 'pragmatic stance' and we look at what perceptual constancy *does*, a unified approach across different modalities seems especially appropriate. Constancy seems to provide perceptual stability for the sake of action in many different modalities, even those that aren't normally understood as being capable to recover specific perceiver-independent properties from an underdetermined proximal stimulus. I take this to be the main point in favor of looking for an alternative to the traditional approach to constancy. Such an account will be spelled out more in the third and fourth chapter. For now, a philosophical desire for systematicity and an appreciation of the ultimately pragmatic goals of perception should be enough, I think, to embark on this journey with me. Whether the destination is worth reaching, only time will tell.

3.0 Perception, structures, and agency

This chapter is divided into two parts. In the first part (sections 3.1. and 3.2.), I connect back to what I have argued in chapters 1 and 2. I want to put forward an account of perceptual constancy which, on the one hand, is genuinely alternative to the traditional one and, on the other hand, relies on different assumptions and approaches the topic from a different angle. Through my proposal, I want to enhance our understanding of constancy as a capacity of perceptual systems but also to elucidate the role it plays in conscious perceptual experience and action. To do this, however, it is necessary to take a step back from constancy and consider perception more generally.

In the first part of this chapter, thus, I discuss the relationship that holds, according to the tradition, between empirical models in perceptual science and the nature of perception. Then, I will argue that some of the worries raised in this context apply to certain not-explicitly-traditional approaches to specific perceptual capacities such as relational color constancy and Isaac's structural realism (Isaac, 2014, 2019a, 2019b, 2019c). I will conclude that, if we want to fully escape the traditional framework and still provide a philosophical account of perceptual constancy consistent with our best scientific models, we should adopt a broadly pragmatist stance.

The second part of the chapter (sections 3.3. and 3.4.) will give a brief historical overview of the broadly pragmatist framework I favor. Through this historical 'detour', I will establish a few 'cornerstones' of the view of constancy which will be articulated and unpacked in chapter 4. Let's begin by presenting what I take to be a problematic view of how perceptual science and philosophical accounts relate to each other.

3.1 The 'Burgean move'

Philosophers within the tradition (e.g. Burge, 2005, 2010; Matthen & Rescorla, 2015) sometimes imply that the success of perceptual science counts as evidence that perceptual systems are literally the way the science represents them to be. In particular, the fact that perceptual psychology is a "serious science" (Burge, 2005, 9) should convince us to take it into account when philosophizing about perception: "There is no reason to doubt that [perceptual psychology] provides insight not only into the mechanics of perception, but into aspects of its nature." (*Ibid.*). As already mentioned in chapter 1, Burge and his followers describe perception as primarily trying to resolve the massive underdetermination afflicting proximal sensory stimuli (e.g. retinal images) in order to reach accurate representations of the distal environment.

On this view, widespread underdetermination is also the main 'obstacle' perceptual psychology encounters when modeling perceptual processes: in addition to the 'natural' underdetermination biological perceptual systems already have to deal with, experimental set-ups and modeling practices add further layers of it. Yet, scientific models are our best chance to gain insight into the nature of perceptual processes. This is because scientists building these models are, at least to an extent, 'in control' of how they deal with underdetermination: they can decide which aspects of perceptual processing to treat as "interfering complications that can, for many explanatory purposes and problems, be idealized away." (Burge, 2005, 12). The fact that these idealized models have "led to a considerable amount of detailed scientific knowledge" (*Ibid.*) is enough, according to Burge, to grant them the power to uncover aspects of the nature of perception. This is the 'Burgean move', which in what follows I will identify with the following claim:

Burgean Move: The empirical success of perceptual science, with all its idealized models and abstractions, is proof that such a science uncovers (aspects of) the 'true nature' of perception and perceptual systems.

The Burgean move relies on two interrelated assumptions. The first is that science (and perceptual science in particular), regardless of its declared goals, can be informative about the nature of the systems and processes it studies. The second is that the widespread abstraction and idealization characterizing scientific models are not only compatible with the fact that science uncovers (aspects of) the true nature of the world, but they are what actually makes it possible. While these assumptions might look prima facie quite innocent, there are good reasons to think that they are false.

A first consideration is that the very definition of idealization in science (e.g. Cartwright, 1983; Potochnik, 2017) seems in tension with the second assumption underlying the Burgean move: "*Idealizations* are assumptions made without regard for whether they are true and often with full knowledge they are false." (Potochnik, 2017, 42). If scientists idealize their models while knowing that such idealizations are false, then it is possible that the very process of idealization substantially affects what a scientific model can and cannot tell us about how the target phenomenon or system a*ctually* is. In particular, if scientists deliberately make false assumptions in the construction of their models, it is probably because they are not interested in truth *per se.* Burge doesn't seem to think that this fact undermines the move, though he should.

Moreover, abstraction and idealization seem to play an essential role in science's actual effectiveness in predicting and intervening on natural phenomena, which, arguably, is a crucial aspect of scientific 'success'. Think about how causally complex the world is compared to our limited representational capacities: without idealizing and simplifying, we would have no way to even begin to understand those aspects of the world that we care about. Yet, so far, through science we have been quite successful at predicting and manipulating the world (even too much, if we consider the level at which human activity has modified the ecosystems on this planet and triggered

climate changes that are now threatening the planet's own capacity to sustain life as we know it!). But precisely because the world is so incredibly complex that it is literally impossible to represent each and every aspect of even just a single phenomenon or system at a time, we should be cautious about attributing to science the goal of accurately describing how the world is. Indeed, if accurate description of systems and phenomena were the goal, even our best models would be still deeply inadequate and, consequently, it would be mysterious how they still managed to be so successful with respect to prediction and manipulation.

Science has goals that are largely *practical*, such as finding ways to exploit relationships and regularities that emerge thanks to, and perhaps *in virtue of*, experimentation and modeling. On this view, idealization and abstraction are themselves what allows science to reach its goals, because they make the world (or at least those aspects of the world that are relevant for human life) 'simple enough' for us to engage with. As Potochnik points out: "the scientific value of simplicity, generality, and the potential for intervention and control are well appreciated" (2017, 42). However, saying that scientific models include abstraction and idealization is *not* the same as saying that they are full-blown fictions. As Chirimuuta (2016) notes, "models are more than tools for predicting future observations—that is, they put us "in touch with" a reality beyond observations and sensory appearances" (755). Idealizations are human-dependent simplifications, but they are also constrained by the world in other respects. And even though sometimes idealization is described as a way of simply 'ignoring' complexity, it goes importantly beyond that: "We can think of models as devices that aim to achieve a certain fit between a natural phenomenon, the human mind, and our particular purposes. Explanatory, predictive, and practical success are a matter of achieving the right kind of fit, not of the attainment of some God's-eye view on the subject" (Chirimuuta, 2016, 756). Therefore, acknowledging that neither scientific

models nor the knowledge they produce are completely human-independent is not the same as claiming that they are fully human-dependent.

That said, however, it is still the case that the abstractions and idealizations contained in scientific models are mostly incompatible with the idea that these models veridically represent or describe the nature of their target system. As already stressed, model-building is aim-relative: no model by itself can be seen as revealing the 'truth' regarding a particular phenomenon, and no model by itself captures all the aspects of a complex process (e.g. Longino, 2013; Mitchell, 2003). Models and experiments have 'local' significance. They are valuable insofar as they are used as the scientists intended, and they are knowledge-producing insofar as one draws inferences using them as tools in a way that's mindful of the underlying theoretical assumptions. [21]

All of this is particularly evident in the context of psychology and cognitive science. Indeed, it has been argued that, due to the very nature of biological cognitive systems (e.g. Godfrey-Smith, 2016; Haugeland, 1996), modeling practices in these areas have basically no choice but to rely heavily on abstraction and idealization (e.g. Chirimuuta, 2019; Godfrey-Smith, 2016). Cognitive systems, like other biological systems, exhibit a degree of complexity that makes 'division of labor' and the simultaneous presence of a plurality of models capturing only parts of the system unavoidable (cfr. Mitchell, 2002). Abstraction and idealization are widely employed in the construction of models of cognitive (and perceptual) systems, and different models bring

[21] Paolo Palmieri (2012) makes this point regarding the way in which psycho-acoustics relying on an idealized and 'atomistic' conception of how external stimuli are processed by the auditory system – namely, by analyzing/decomposing the sound stimulus into its Fourier components and then 're-building' a more simplified combination of the single sinusoidal signals: "Analysis is always relative to a system of representation. It is useful insofar as the system of representation serves some practical purposes; for instance, in the case of pure tones and Fourier analysis, that of making telephony possible. It becomes a philosophical hindrance when it is hypostatised into an ontology of fundamental entities to be discovered in the world." (Palmieri 2012, 541).

different ‘perspectives’ on the system at hand, without necessarily being incompatible, at least at the theoretical level. In fact, this kind of perspectivism, it has been argued, might be the *only* reasonable way of understanding scientific practice and model building in this field (Chirimuuta, 2018; Giere, 2006). Because of this, it is once again hard to envision how the Burgean move could be granted, especially in the case of perceptual science, where each perspectival model only offers, at most, a ‘partial truth’ about the nature of the relevant system.

We saw an example of how multiple perspectives can co-exist, at least at the theoretical level, in the case of color constancy (section 1.4 in this document). Color constancy research is a good example of a scientific domain in which we have two models that could be described as equally empirically adequate and yet having different (though perhaps not incompatible) theoretical commitments. In light of this, the most reasonable conclusion to draw is that neither model discloses the ‘true nature’ of color constancy, and not for lack of trying. It is simply not what a single model of such a complex system can do on its own. Once again, this is in tension with the Burgean move. I will now move to examining a different, less traditional approach, and argue that it is nonetheless vexed by problems similar to those affecting the Burgean move.

3.2 Isaac’s ‘fixed points realism’ and the nature of perceptual systems

Even though scientific models do not fully disclose the ‘true nature’ of the systems and phenomena they target, which is what created trouble for the ‘original’ Burgean move, perhaps our best scientific models uncover at least *some* aspects of the human-independent world. In

particular, consider the suggestion made by Alistair Isaac (2014, 2019a, 2019b, 2019c). Isaac endorses what he calls Fixed Points Realism (FPR), and grounds it in the notion of *measurement*:[22]

> Values obtained through successful measurement veridically represent objective *fixed points* in the world, which may be uniquely specified by the pattern of distances that obtain between them in a metric space. […]
>
> By "fixed point," I mean a modally stable phenomenon, a feature of the world that reveals itself across a variety of possible interventions. Such points are fixed independent of our interests or theoretical commitments: *they are discovered, not constructed*. The objectivity and theory-neutrality of these points follows from their susceptibility to successful measurement: convergence on some value eliminates the possibility that any particular theory plays a constitutive role in its determination; increased precision requires the cooperation of nature, thus revealing objective fixity in the phenomenon that value represents. (9)

Many mathematical equations can represent the relations existing among fixed points, and nothing about fixed points commits Isaac to particular entities occupying those points in geometrical space. All that the discovery of fixed points tells us, according to Isaac, is that there are relations in the world that can be represented by geometrical relations among points in space, and that our best measurement practices and techniques allow us to know them. Fixed points realism, indeed, is a form of *structural* realism.

Structural realism (e.g. Ladyman, 1998; Massimi, 2011; Worrall, 1989) is the view in philosophy of science according to which successful scientific inquiry manages to uncover (or discover) a system of relations, or structures, that exists in the world, independently of any scientific theory or hypothesis. These structures are usually thought of as being 'made explicit' by our best mathematical equations and physical laws. According to the structural realist, and in particular a structural realist of the *epistemic* kind like Isaac, science does not have the 'power' to

[22] "I take *measurement* to be any empirical procedure for assigning points (or regions) in a metric space to states of the world, where a *metric space* is any set of elements with a distance metric defined over it." (Isaac, 2019a, 5).

discover exactly how the world is in all of its aspects.[23] For example, it might turn out that all the currently unobservable entities that scientific theories postulate and treat as 'real' do not really exist independently of the theory that postulates them. We cannot be 'entity realists' because we, as scientists, do not have access to the most fundamental components of the world. What science *can* do, however, is find out how these fundamental components (if there are any at all) relate to each other. Relations can be discovered, studied, and observed even in a case in which what we think the *relata* are is completely wrong. Our best science, according to structural realists, allows us to have access to 'how the world really is' as long as we interpret the expression as referring to observation- and theory-independent structures and relations as they are represented in measurements, and not to the entities standing in such relations.

In fact, Isaac's version of structural realism is very modest. He doesn't commit himself to realism regarding structures like "the rich mathematical structure of physical theory" (2019a, 9). All Isaac is willing to be realist about is fixed points in geometric space and their relative positions:

> Since measurement maps the outcome of an empirical procedure into a metric space, it also provides a theory-neutral "language" for describing fixed points, namely their relative positions within this space. While our theory of temperature may change radically, and with it even our theoretical terms for key values, such as those we now call "boiling point of water" and "absolute zero," the relative positions of these values cannot change. (Isaac, 2019a, 9)

This view of scientific practice takes into account the role of idealization and abstraction in the construction of models and acknowledges their 'perspectivalness'. Even though this modest form of realism keeps Isaac away from (a non-perceptual-science-specific formulation of) the Burgean move, I will argue that applying FPR to perceptual psychology generates some

[23] We can distinguish two kinds of structural realism. 'Epistemic' structural realism (e.g. Worrall and Isaac) is the view according to which all we can *know* about the ultimate constituents of reality is that they stand in certain structures and relations. On the other hand, 'ontic' structural realism (e.g. Ladyman) claims that structures themselves are the ultimate constituents of reality.

problematic commitments for Isaac, which ultimately put him in no better position than the defender of the Burgean move.

In his (2014), Isaac applies an early version of FPR to psychology, and in particular to the branch of psychology that deals with the relationship between our conscious sensations and their external causes. Isaac argues that our sensations, for instance, of heat, are the relevant measurement outcomes in the specific context, just like thermometric measurements are outcomes for physics:

> I claim that the sensation of heat stands in the same relationship to temperature as the outcome of a simple thermometric measurement. Sensations of heat are linearly ordered against a neutral baseline. Above this baseline, we describe them as more or less warm, below the baseline, more or less cold. Just as a single thermometer can be calibrated differently to deliver numbers on either the Fahrenheit or Celsius scale, so also the same physiological apparatus may be calibrated against different baselines when generating sensations of heat. (2014, 8)

A similar story is repeated for color:

> the relationship between color sensations and properties in the world is analogous to the relationship between a measuring space and a measured space. […] Our possible experiences of color are organized into a geometrical space commonly called the color solid. […] The color solid characterizes the relative distances between possible experiences of color, as determined through assessments of color similarity. (2014, 10)

But if sensations are measurement outcomes, and measurement of sensations can be considered successful, then – and this simply follows from the definition of FPR – measuring sensations and 'ordering' them on a certain independent scale discloses fixed points and their mutual relations in the world: "If we accept this analogy, then it appears that the relationship between heat [or color] as *we experience it* and heat [or color] *as it is in the world* is one of structural correspondence" (Isaac, 2014, 9; emphasis in the original text).

In light of all this, I think that we can reconstruct the following argument out of Isaac's view applied to perceptual psychology, even though he does not endorse it explicitly. This argument leads to a conclusion which is not only controversial, but likely false. If my

reconstruction is right and Isaac is in fact committed to this conclusion, then we have a reason to reject FPR in the context of perceptual psychology. The argument goes as follows:

P1: According to FPR, successful measurement outcomes in science veridically represent relations among 'fixed points' in the objective world. (from Isaac, 2019a)

P2: Perceptual psychology is a science.

P3: The successful measurement outcomes in perceptual psychology veridically represent 'fixed points' in the objective world. (from P1 & P2)

P4: Sensations veridically represent relations among 'fixed points' in the objective world (from Isaac, 2014, quoted above)

P5: Measuring sensations is part of the practice of psychophysics, which in turn is part of perceptual psychology.

P6: Sensations in perceptual psychology are analogous to measurement outcomes in other sciences. (from P3, P4 & P5)

P7: If sensations in perceptual psychology are analogous to measurement outcomes, then sensory systems are analogous to measuring instruments.

Conclusion: Therefore, according to FPR, sensory systems are analogous to measuring instruments. (from P6 & P7)

The premise that needs a more elaborate defense is P7: I'll try to provide one now. According to Isaac, sensations are analogous to a measuring space, while the properties in the world such sensations represent are analogous to the measured space (2014, 10). Measuring means mapping the measured space into the measuring space, or, in other words, successfully capturing the isomorphism between the structure of the measuring space and the structure of the measured space. The measuring space is set by the measuring instrument: instruments are the 'interface' between measuring space and measured space. Continuing with the same analogy, properties in the world are measured and the outcome of such measurements are sensations. Thus, if sensations map onto/measure properties in the world, then sensory systems must be those 'making' the measurement, and P7 is granted. The conclusion, then, states that the sensory systems of living

organisms are analogous to measuring instruments: sensory systems just are for perceptual psychology what thermometers are for physics.

But what would it be for a sensory system to be analogous to a measuring instrument? Measuring instruments detect magnitudes and represent the relations among different values of such magnitudes on a pre-determined scale. Therefore, the claim that perceptual systems are analogous to measuring instruments implies that there must be something 'out there' that those systems, when working successfully, detect, and to which the categories we employ in making sense of our experience must *correspond to*. Indeed, that sensory systems have the function of detecting and measuring is a pretty old assumption in philosophy. However, this by no means entails that it is correct. If we look at how perceptual scientists understand their own discipline, none of these ideas seem to be explicitly endorsed. Indeed, we have good reasons to think that perceptual science as a whole is more sympathetic to a theory of the function of perceptual systems called "perceptual pragmatism" (Chirimuuta, 2015, 2017):

> Consider the question "what are perceptual states for?" A natural response is "to tell us what is where." This chimes easily with the correspondence way of thinking, but it is not the most fundamental of responses. The answer that predominates in perceptual science is "to help you to live by guiding your activity in the world," and "telling you what is where" *is but one way* to achieve this. (Chirimuuta 2015, 110; emphasis mine).

If perceptual systems are 'pragmatist' in this way, then we have a reason to reject the idea that they might be detecting any perceiver-independent element in the world at all, not even very 'basic' ones such as Isaac's fixed points. If anything, models like the "color solid" (Isaac 2014, 10) or its analogs in other sense modalities (e.g. Shepard's "toroidal pitch space", also mentioned in Isaac, 2014, 24) are considered useful instruments to explore interesting regularities in how we experience the world. But the fact that regularities in our experiences can be identified, studied, and even predicted should not be seen as proof that, through measuring our sensations and mapping

them onto a metric chart, we are in fact tracking perceiver-independent structures. Some perceptual psychologists even argued that perceptual systems evolved systematically to represent the world inaccurately, because certain distortions actually help creatures in their interactions with the environment, and therefore maximize the 'fitness' of the creature (e.g. Hoffman, 2016, 2018; Hoffman & Singh, 2012; Mark et al., 2010).

Whether systematic perceptual misrepresentation is actually the case or simply another idealization, one thing seems clear: it seems wrong to assume without further justification that perception veridically representing *anything* perceiver-independent in the external world. In turn, this suggests that the analogy between sensory systems and measuring instruments should be rejected. This finally delivers the conclusion that even the 'modest' form of realism defended by Isaac, which he applies to perceptual psychology by means of the analogy between sensory systems and measuring instruments, does not correctly describe what perceptual psychology does and what it cares about. The Burgean move, stronger or weaker in its formulation, is nonetheless always in the background of any attempt to take a realist stance on perceptual science. Because of this, no realism with respect to perceptual science can succeed: if our perceptual systems are 'pragmatist', so should we. The historical detour in the next section will provide some context and key notions to start building my non-traditional view of constancy, which is indeed pragmatist in spirit.

3.3 Historical detour: all roads lead to William James

Both James (1890, 1909, 1912) and Dewey (1896) share one main background assumption which separates their view of perceptual psychology from the traditional one. This assumption is

that perception and action are essentially 'for one another', that is, perception and action are interdependent and part of the same whole (which Dewey calls "coordination"). In this picture, perception cannot be fully understood or accurately described unless we look at how our perceptually guided activities unfold over time. This can be done best by observing and studying how a creature gets by in her 'natural' environment, rather than by designing 'artificial' tasks in the lab which already make assumptions about what kind of capacities will be needed in order to perform them.

This pragmatist view denies that we can understand how perception really works by only studying 'crystallized' moments of what is in fact a continuous activity of exploration and a constant 'stream' of sensory stimulation coming from all sides and through all channels. Sometimes, this crystallized view is also called the 'snapshot' conception of perception (e.g. Noë 2004), and it usually goes hand in hand with what Chirimuuta (2015) calls the "detection model", which informs the analogy between perceptual systems and measuring instruments I rejected in the previous sections. According to Chirimuuta, the alternative to the detection model is the "utility model", which has at its core the idea that perceptual systems have evolved with the main function of 'telling' a creature what's useful for her survival, whether or not this includes accurate details regarding how the world actually is. Sometimes, and this is the main insight of the utility model, what's useful doesn't correspond to what's 'true' independently of the perceiver. In turn, this idea is more compatible with seeing perception as an activity whose main goal is to guide action, because, for creatures like us, action is fundamental to survival.

Think, for example, about color vision. Why did humans evolve a visual system that allows them to see in color at all? If one endorses the detection model, one will answer that, because colors are a way to pick out actual properties of surfaces (SSRs or similar), color vision allows us

to see the world in a more accurate way, by presenting us with details that, for instance, a dog wouldn't notice. On the other hand, if one prefers the utility model, one will say that color vision doesn't have representational accuracy as its main goal. Rather, we see in color mostly because it's useful, for creatures like us, to have ways to discriminate visually things that are good for us from things that aren't, and to guide our actions accordingly. We can, for example, discriminate ripe fruit from leaves by relying on their color. A dog might not need color vision in the same way, because it relies more on olfaction when it comes to discriminating food from non-edible stuff.

The approach to perception that lies behind the utility model starts developing in a systematic way in the mid-nineteenth century. At that time, Charles Darwin (1859) and his model of evolution by natural selection was becoming more and more famous among biologists. As a consequence, terms like 'adaptation' start circulating among psychologists and attracting their curiosity. Among the psychologists most impressed by the new evolutionary theories being developed there is William James. James was indeed inspired by Darwin – as well as by Herbert Spencer's work in both biology (1864-67) and psychology (1855/1870) – when he argued that the science of psychology should adopt a broader conception of cognition as embedded in an environment and intrinsically influenced by it. James thought that psychology should study cognition at the 'functional level' (Gibson will call it the 'ecological level'), that is, the level at which a living organism and her environment are inseparable, and one cannot be understood independently of the other. Doing psychology this way was meant to be in clear contrast with the psychophysical study of perception based on measuring the intensity of sensations and analyzing them separately from perceptual 'constructions', which was the preferred methodology of first-generation Germany-trained psychologists such as Wundt and Titchener (Chirimuuta 2016).

Our cognitive activity, James argued, is a continuous attempt to *adapt* to the challenges of the environment, just like other physical features evolved as adaptations to the external circumstances in which a creature had to survive and reproduce. Having a cognitive life, in this framework, is just one among many adaptations that have made creatures like us fit for survival. It is emblematic that, at the very beginning of *The Principles of Psychology* (1890, 19) he wrote that "few recent formulas have done more real service" than this claim by Spencer: "the essence of mental life and bodily life are one, namely, 'the adjustment of inner to outer relations'". Even single mental states can be seen, according to James, as a 'temporary' adaptation to our current surroundings. For example, my visual experience of a tree is the result of a temporary adaptation of my visual system to the pattern of light reflected by the tree in front of me. As my circumstances change, my mental states change, too, as they adapt to the former with no interruption. Mental activity, in this picture, should not be 'broken down' into single experiences of single properties. Rather, it should be studied as a network, where the relations among different states are more fundamental than their individual contents. Similarly, what counts as 'inside' ad what counts as 'outside' in the context of this adaptation-based view is only arbitrarily defined, as the adaptation can go both ways: from perception to action, and from action to perception. Indeed, James (especially in later works) and Dewey are among the fiercest opponents of the stimulus-response distinction, which grounds most experimental psychology today and has done so since the rise of behaviorism in the 1920s. For them, the very idea of distinguishing an 'input' (stimulus) phase from an 'output' (response) phase in the analysis of how cognition works misunderstands the nature of cognition itself, and it is, essentially, a way to re-introduce Cartesian dualism.

Because cognition is, in the pragmatist picture, a constant dynamical exchange of information within a perceiver-environment unified system, it doesn't make sense to try and draw

the line between what happens ‘in the mind’ and what happens ‘in the world’, or, indeed, between what’s considered a ‘perceptual input’ and what’s identified as a (generally motor) ‘output’. The identity and the significance of each cognitive process is determined by the context, and by the rest of the cognitive network the process is embedded in. Dewey (1896) provides a beautifully described example. Think about experiencing a loud, unexpected sound:

> If one is reading a book, if one is hunting, if one is watching in a dark place on a lonely night, if one is performing a chemical experiment, in each case, the noise has a very different psychical value; it is a different experience. In any case, what proceeds the ‘stimulus’ is a whole act, a sensori-motor coordination. What is more to the point, the ' stimulus' emerges out of this coordination; it is born from it as its matrix; it represents as it were an escape from it. (361)
>
> And if your ‘response’ is to run away…
>
> the final ‘element’, the running away, is not merely motor, but is sensori-motor, having its sensory value and its muscular mechanism. It is also a coordination. And, finally, this sensori-motor coordination is not a new act, supervening upon what preceded. Just as the 'response' is necessary to constitute the stimulus, to determine it as sound and as this kind of sound, of wild beast or robber, so the sound experience must persist as a value in the running, to keep it up, to control it. The motor reaction involved in the running is, once more, into, not merely to, the sound. It occurs to change the sound, to get rid of it. (363)

Later on, J. J. Gibson makes a somehow analogous point when he claims that “animal and environment make an inseparable pair. Each term implies the other.” (1979, 8). This is particularly true when we think about what it is to behave in an environment, responding to certain perceived features and having our movements guided by them. Take, for example, the act of reaching and grasping:

> Whereas it is possible to talk abstractly about reaching and grasping as pure acts, independent of an environmental referent, in fact their *psychological expression* is always based on the co-influence of bodily and environmental properties. […] An individual does not reach per se, but reaches toward something; and, importantly, the biomechanics of reaching vary as the relative location of the object to the body changes. (Heft 2001, 110).

However, Gibson takes and makes his own only part of the whole picture behind the pragmatist ideas of 'coordination' and 'adaptation'. In particular, Gibson didn't fully embrace the late James's focus on embodiment and the stress on purposes as actively shaping perceptual experience even of features which are not immediately connected with movement and behavior. For Gibson, a creature's body is a moving entity in space rather than the 'center of the universe' in a heavier, metaphysical sense, which includes non-spatial features of the embodied perceiver like purposes and habits. In the pragmatist framework, however, these 'intentional' features of perceivers are fundamental and ineliminable if one wants to understand perception in all its complexity, which, indeed, includes its relationship with action and the sense in which it grounds objectivity. Instead, there are striking similarities with the late James's 'existentialist turn' in the work of Merleau-Ponty (1963, 1964, 2013). This passage, a footnote from James's *Essays in radical empiricism* (published posthumously in 1912), could belong in one Merleau-Ponty's works:

> The world experienced (otherwise called "the field of consciousness") comes at all times with our body at its centre, centre of vision, centre of action, centre of interest. […] The body is the storm centre, the origin of co-ordinates, the constant place of stress in all that experience-train. Everything circles around it, and is felt from its point of view. (1912, 86).

Merleau-Ponty's phenomenological account of perceptual constancy and the objectivity afforded by perception[24], on which I draw in developing my positive view, have James's radical empiricism as one of their main precursors. A discussion of how Merleau-Ponty informs my own proposal regarding perceptual constancy and objectivity, however, will have to wait until the next chapter. Now it is time to take stock and apply the pragmatist points discussed in this section to back to perceptual constancy.

---

[24] See also Kelly (2005, 2010) and Matherne (2017).

3.4 Purposeful action as the 'missing ingredient'

I want to borrow from James and Dewey essentially the idea that perception and action are ultimately one and the same thing[25], and we cannot understand the former without looking at the latter. For this reason, my proposal consists in incorporating action *within* an account of constancy, which is a capacity involved in all sorts of perception-guided cognitive processes, and not just as what constancy, among other things, causally enables. If action and perception constitutively belong together as James and Dewey argued, then perceptual constancy, which is a perceptual capacity, is more closely tied to action than it would be within a traditional framework, where perception enables action without being itself part of it.

Consider this fact: arguably, we developed the capacity for constancy – and evolved the sub-personal mechanisms able to realize it – in order to solve some kind of 'problem'. This problem is that, when we move, our sensory receptors are capable of 'taking in' an overwhelming amount of information. So much, in fact, that our brain couldn't possibly process it all *and* make our body move around at the same time. Imagine having to understand what someone on the radio is saying in order to act on the basis of it while a huge amount of frequencies are all audible and interfere with the only one you're interested in. This is why we needed a 'psychological corrector', something that allowed only *some* information to be used to coordinate our movements towards particular targets. If we didn't move, the information 'coming in' at any given moment in time would be much less, and we wouldn't need a way to stabilize the perceptual scene in the same way. And when do we usually need or want to move? When we need or want to *act*. If we didn't need to act in order to survive, we wouldn't move. We would be mostly inert, like sea sponges.

[25] This idea also constitutes the 'heart and soul' of Gibson's ecological psychology.

Without goal-directed movement, constancy wouldn't really need to exist, or at least not in the pervasive and sophisticated way in which it is present in humans.

Now, consider that, if constancy has to help us act, and at the same time it did so by trying to recover the 'absolute' perceiver-independent properties of objects, acting would require a good amount of preliminary reflection in order to figure out how to coordinate my movements with a perceived scene that is tending towards a 'view from nowhere', as perspectival information would be discounted (more on the apparently problematic relationship between movement and the traditional account in chapter 4, section 4.3).

But the way in which perception guides action isn't, in the most basic sense, mediated by reflection: I don't need to think in order for my hand to adjust its grip to the shape and size of the cup I am reaching towards. In contrast, it seems to me that constancy can play its role only if variance and invariance are *both* present and actively playing a role in proximal stimuli, in particular by standing in precise relations with each other. In the past, people sympathetic with pragmatism have already suggested to move away from the traditional conception of constancy – i.e. the conception of constancy that focuses on the recovery of specific perceiver-independent properties. Here is, for example, Ittelson (1951):

> The most serious limitation of such measures of correspondence [Ittelson is talking about the experimental approach to constancy that aims at 'measuring' how accurate our attribution of properties to object is under different conditions] is that they channel interest away from the means by which the organism achieves constancy toward a description of the final achievement. […] Constancy mechanisms and constancy achievements are inseparable. Any complete theory of perceptual constancy must encompass all its aspects. (286)

Ittelson was part of a movement called 'transactional functionalism' (Noble 1981), which, at the beginning of the 1950s, was trying to make mainstream behaviorism more sensitive to pragmatist and Gestalt psychology insights and methods. Ittelson and other transactional

functionalists (e.g. Kilpatrick 1952, Ames 1953) tried to 'import' from pragmatist philosophy, for example, the idea that perception is 'transactional' (Dewey 1929), that is, that perception should be understood as a continuous 'exchange' between the perceiver and the world. The world is 'changed' by the fact that a perceiver experiences it and acts in it on the basis of her experience, and the perceiver is 'changed' in that her goals and actions 'shape' her way of perceiving the world. Action and purpose lie at the core of a creature's perceptual capacities, and one cannot understand perception without understanding its relationship with action. Here is another passage from Dewey (1896):

> what is wanted is that sensory stimulus, central connections and motor responses shall be viewed, not as separate and complete entities in themselves, but as divisions of labor, functioning factors, within the single concrete whole, now designated the reflex arc. (358)
>
> the reflex arc idea, as commonly employed, is defective in that it assumes sensory stimulus and motor response as distinct psychical existences, while in reality they are always inside a coordination and have their significance purely from the part played in maintaining or reconstituting the coordination; (360)

Treating proximal stimulation and the distal, perceiver-independent properties of objects as two different elements, or 'steps' in order to study perceptual constancy is, for this reason, a way of misunderstanding what constancy primarily *does,* what it's *for:*

> Perception, then, is the product of the continual recording of the relatedness of things *as defined by action*. Perception is the apprehension of significance. […] Such an approach makes possible a more adequate understanding of the perceptual constancies. (Ittelson 1951, 290, emphasis mine).

To really do justice to the complexity of the proximal-distal relation, we have to include action and agency in our account of what kind of capacity perceptual constancy is. Even though Dewey didn't explicitly talk about constancy, I think his observations concerning the distinction between 'stimulus' and 'response' are very much in line with what I would say in response to the tradition. If by 'stimulus' we mean what the tradition says a display of constancy is, namely

proximal sensation + distal property 'attribution', and by 'response' we mean what I argued they have mistakenly left out, namely action, the following passage is basically a summary of what I suggest with respect to constancy: "The fact is that stimulus and response are not distinctions of existence, but teleological distinctions, that is, distinctions of function, or part played, *with reference to reaching or maintaining an end*" (Dewey 1896, 365; emphasis mine).

In the same way, I claim, perceptual constancy is a capacity that constitutively involves action, and its main goal is to favor sensorimotor coordination and active engagement with the world. The tradition treated perceptual constancy as 'merely' enabling action, instead of seeing action as one of the very constituents without which the capacity cannot be fully understood. Description of capacities and accounts of phenomena are usually backed up by theoretical assumptions that are, most of the time, not innocuous, and can be challenged if we have good reasons to do so.

Not incorporating action into a view of constancy has been the main mistake of the tradition. A mistake that was diagnosed by transactional functionalists like Ittelson and that can be found, to an extent, in relational, ecological, or structural realist accounts, too. Indeed, I think that leaving agency out comes at the price of losing sight of what, especially according to philosophers like Burge, made constancy such an intriguing and significant perceptual capacity: namely, its potential to be the key to explain perceptual objectivity, that is, perception's ability to present the world as mind-independent. The relation between constancy and objectivity is discussed in the next (and last) chapter.

4.0 The Relational Invariance view and objectivity

4.1 Introduction

Throughout the previous chapters, I stressed the idea that, to completely understand perceptual constancy, we must understand its role in providing the perceptual invariance needed for action. I have argued that the capacity to represent objects' perceiver-independent properties is quite distinct from the capacity to act successfully on the basis of perception. This is mostly due to the fact that there are cases in which the two capacities can come apart, like in the case of D.F. and her condition called 'visual object form agnosia' (Goodale & Milner, 2013; Milner & Goodale, 2008; Milner & Goodale, 1995), thus supporting the idea that they are functionally distinct and that, for this reason, the former cannot be used to explain the latter. Given that, on the one hand, people in the tradition have identified constancy with the capacity to represent perceiver-independent properties but that, on the other hand, it seems implausible to say that one could act successfully on the basis of perception without possessing the capacity for perceptual constancy, I have concluded that the traditional view should be put aside. If we want to do justice to the role of perceptual constancy in perceptually guided action, we should adopt a different view, which I am presenting next.

On this view, perceptual constancy is the capacity of perceptual systems (visual and not) to keep track[26] of invariant relations in the environment (which is *not* the perceiver-independent physical world) specified by lawful patterns of (sometimes perceiver-generated) variation in

[26] One can replace 'keep track' with 'represent' at will, as my account is neutral with respect to whether perception is a representational capacity or not.

proximal stimuli over time. First, I will discuss once again the issue of constancy in action and go through the argument for why the tradition cannot account for it. Secondly, I will present the 'Relational invariance view', together with some implications and potential objections. Finally, I will connect the view of constancy to another notion that philosophers of perception hold dear: the *objectivity* of perception.

In particular, Burge (2010) thinks that perceptual objectivity is grounded in and depends directly on constancy. Even though I have argued that Burge's favorite view of constancy is, if not completely wrong, at least importantly incomplete, I still agree with him that constancy is key if we want to understand what it is to perceive objectively. In this chapter I am going to argue that, if we want to accept Burge's claim that constancy is where objectivity 'comes from' while at the same time taking on board my view of constancy, then our understanding of perceptual objectivity itself must change. And how should it change? I will answer this question by presenting a view of perceptual objectivity that differs from Burge's, and is consistent with my view of constancy: this alternative view is mostly due to Maurice Merleau-Ponty (Merleau-Ponty, 2013; 1968). I will conclude that Merleau-Ponty's idea of perceptual objectivity and the Relational Invariance view of perceptual constancy together form a new framework that has the potential to explain a lot about the nature of perception and its relation to action.

4.2 Constancy in action: a problem for the tradition

As already discussed in chapter 1, it is widely accepted, at least in vision science, that the capacity to represent perceiver-independent, categorical properties (and to attribute them to specific objects in judgment) and the capacity to rely on perception to guide motor action can be

possessed independently of one another (Milner & Goodale, 1995, 2008; Goodale & Milner, 2013; Mishkin & Ungerleider, 1982). Indeed, the two capacities are realized in two different parts of the visual cortex, specifically in two different neural 'pathways'. The 'ventral' pathway (or stream) works with 'allocentric' (which, for our purposes, can be equated to 'perceiver-independent') representations of objects and their properties, and enables conscious attributions of these properties in judgment. On the other hand, the 'dorsal' pathway (or stream) represents 'egocentrically', that is, works with representations that take into account the perceiver and the way in which *she* relates to objects (e.g. Deubel et al., 2010; Helmholtz, 1867; Wu, 2014). Visually guided motor action depends on the activity of the dorsal stream, and relies on egocentric, and thus perceiver-dependent visual representations. Wu (2014, p. 385 and elsewhere) calls the capacity of the dorsal stream to stabilize sensory inputs so that egocentric relations with the world are represented and can guide action "spatial constancy", and he uses it to explain the fact that, when we move our eyes to look around a room, we do not experience the room itself and the objects in it as changing position, even though the position of their projection on the retina changes.

The main source of evidence for the (functional and anatomical) mutual independence of the capacity to represent perceiver-independent properties and the capacity to rely on perception to guide motor action is Milner and Goodale's study of D.F. (see section 1.3 in this document). Recall three main theses that make up the traditional view of constancy:

Perceptual Invariance: Perceptual constancies are the capacities of perceptual systems responsible for the kind of perceived invariance *necessary* for stable conscious perceptual experience, accurate perceptual judgments, and successful perceptually guided motor actions.

Traditional Constancy Thesis: Perceptual constancy is the capacity of perceptual systems to recover the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli.

Constancy in Action: Recovering the perceiver-independent properties of objects by discounting contextual and perceiver-dependent elements in proximal stimuli is necessary for successful perceptually guided motor action.

The case of D.F. constitutes a counterexample to *Constancy in Action* and, consequently, to *Traditional Constancy* (provided that one still finds *Perceptual Invariance* plausible). Indeed, D.F.'s visual system seems to, on the one hand, lack the capacity to represent perceiver-independent properties of distal objects (due to ventral stream damage), but, on the other hand, maintain the capacity to perform visually guided motor actions. Now, the former capacity is what the tradition identifies as perceptual constancy. But if we think that the latter capacity, which D.F. possesses, requires the exercise of perceptual constancy (as *Perceptual Invariance* states), then perceptual constancy cannot be what the tradition says it is.

At this point, there are different routes one could take. One could insist that D.F.'s actions, though visually guided, do not require constancy, which basically is equivalent to rejecting (part of) the *Perceptual Invariance* thesis. This reply is especially tempting if one takes the fact that we often consciously *experience* the perceptual stability enabling actions and judgments as suggesting that constancy itself, being the capacity that enables and explains perceptual stability, must be consciously experienced in order to be successfully displayed. If one considers perceptual constancy to be a perceptual capacity whose successful exercise necessarily involves consciously experiencing the exercise, then it is at least plausible to claim that D.F. does *not* have the capacity for constancy. This is mainly because D.F., by her own reports, does *not* consciously experience a

stable world, but only 'floating' colors and textures.[27] If one chooses this route, one can resist the D.F. counterexample and still to an extent salvage *Perceptual Stability* by saying that, if constancy plays a role in perceptually guided action at all, this is only true of perceivers with no relevant brain damage. Consequently, D.F.'s perceptually guided motor actions are performed in the absence of perceptual constancy.

However, what would perceptually guided motor actions without constancy look like? Would they still be actions? Would they count as perceptually guided at all? Think about the basic structure of a perceptually guided motor action that D.F. can perform, for example grasping a bottle. The motor system receives an input from the visual system, which 'tells' the motor system, minimally, something about how the world is *in relation to* the perceiver's body (since the ultimate goal is to move such body and act on the world). In particular, it conveys information about the spatial relations between perceived and object. On the basis of this perceptual input, the motor system 'decides' which movements of the arm, hand, etc. to produce in order for the bottle to be successfully grasped. Motor actions are extremely complicated coordination problems, where perceptual input and motor output are coupled in a very precise way, and for the most part all of this happens below the threshold of consciousness.

---

[27] It should be noted, however, that the relationship between consciousness and the dorsal stream has been object of some controversy. People disagree about whether the activity of the dorsal stream contributes to conscious experience, whether it should be considered as 'evidence' of unconscious vision, or whether we should just 'downgrade' it to the activity of a sub-personal mechanism (cfr. Peters, Kentridge, Phillips, & Block, 2017). Personally, I find the third option phenomenologically implausible: visually guided actions don't feel as if they are effects of something sub-personal (the way in which, for example, digestion feels). When we act, and arguably when D.F. acts, too, it seems to be something that *we*, the agents, do, not something a certain part of my body does under the 'command' of the brain. Purposefully grasping a bottle is different from having a muscle twitch in my arm so that somehow my hand ends up in the same grasping position. The former is clearly visually guided *and* attributable to a person (i.e. an agent), the latter isn't. Regarding the second alternative, I agree with Phillips (against Block) that the idea of unconscious perception is in itself problematic. The first option seems to me the only one that has some traction, even though the particular case of D.F. does not provide conclusive evidence in either direction.

Now, think about what the perceptual input would look like if percepts weren't stabilized, that is, if constancy mechanisms weren't involved. In visual perception, retinal images would change every few milliseconds, due to saccadic movements of the eye. Because by hypothesis we are in a scenario without constancy mechanisms, the perceptual input for the motor action would consist in these fluctuating proximal stimuli. If the motor system had to re-adjust the motor commands each time the perceptual input changes, i.e. every time we saccade, no motor command would ever be executed, because by the time the signal reaches the peripheral nervous system, the initial command would have to be updated in response to a new perceptual input coming in (corresponding to a change in the proximal stimulus). We would be, in other words, constantly undecided about how to move our bodies, so that we would end up not moving at all. But this is definitely *not* what happens: D.F. moves, grasps bottles, catches balls, even though she does not consciously experience the kind of stability we normally experience when grasping bottles or catching balls.

Therefore, if we accept the claim that constancy is the capacity thanks to which perceptual systems operate the sometimes quite sophisticated 'psychological corrections' of proximal stimuli that make possible , in different ways, (i) the conscious experience of invariance, (ii) judgments, and (iii) perceptually guided actions, then we have to accept that D.F. has constancy. Following this second route, then, one could say that perceptually guided motor action *does* require perceptual constancy and that, moreover, *precisely for this reason* perceptual constancy is not the capacity to recover the perceiver-independent properties of objects.

Regarding the issue of conscious experience, if one takes this second route, one is committed to the idea that, despite what we might be pre-theoretically inclined to think, experiencing invariance isn't itself part of what successful exercises of constancy consist in, but

rather a consequence of it *only occurring* under certain circumstances. On this view, D.F. would be the example of perceptual constancy being successfully exercised in the absence of conscious experience of invariance. This second route takes us to the Relational Invariance view, which is indeed an alternative to the traditional view: it rejects *Traditional Constancy* while preserving *Perceptual Invariance* in its entirety, and especially the bit regarding perceptually guided motor action.

Therefore, the first important characteristic of the Relational Invariance view is that it does not consider consciously experiencing invariance as necessarily involved in successful constancy displays, though such an experience is often associated with the displays, and provides a reliable cue for them. For this reason, in articulating the Relational Invariance view, I will focus on ordinary perceivers, for whom a conscious experience of invariance is associated with constancy displays. A second characteristic is that it does not identify constancy with the recovery of perceiver-independent properties of specific distal objects such as shapes, surface spectral reflectance properties, sizes, sound intensity etc. This last characteristic makes it genuinely alternative to the traditional view. Now it is time to present the Relational Invariance view in more detail.

The Relational Invariance view (from now on, RI) states that perceptual invariance is explained by the fact that certain *relations* within the perceived scene remain unchanged (from the perceiver's perspective) and are therefore represented by the perceptual system as 'global' invariant features in the environment. Perceptual experience of these relational invariants, then, is explained by the capacity of perceptual systems to track regularities in how proximal stimuli change over time, where these regularities specify, at the level of perceptual content, global relational invariants in the scene. Nowadays, RI is favored mostly by philosophers and

psychologists working in the embodied/ecological framework (e.g. Chemero, 2009), but it has its roots in the works of French phenomenologist Maurice Merleau-Ponty (1963, 1964, 2013), as well as American pragmatists William James (1890) and John Dewey (1896).

In this section, I present and defend RI by means of two claims. First, a vast range of invariance experiences (in various sense modalities) seem to have a common and precise role in our cognitive lives: they enable a variety of recognition and re-identification tasks. Second, the proposal that perceptual systems represent relational invariants instead of absolute properties fits better with and better explains this role of invariance experiences. As an alternative to the Traditional Constancy Thesis, RI proposes:

Relational Invariance Thesis: Perceptual constancy is the capacity of perceptual systems to track invariant patterns of relations (sometimes perceiver-involving) among elements in the perceptual scene over time and across changes affecting proximal stimuli.

4.3 The Relational Invariance view

I claim that invariance experiences, i.e. the experiences of certain elements in a perceived scene as remaining unchanged while other elements change, have one main function: they allow us to recognize and re-identify objects, places, events, in different contexts and sets of non-ideal circumstances. Indeed, for us to be able to recognize objects across contexts, there must be *something* among what's being perceptually experienced that remains the same while something else changes, and this is precisely what invariance experiences consist in. Note that sometimes perceptual recognition across contexts is not immediate. Sometimes, it can take us a while to

attribute a certain property to a certain object (e.g. to determine the color of a shirt, or the shape of a table) the first time we experience it – which is presumably a necessary condition for future recognition. Here are a few examples which should suffice in order to see the overall point. For the first example, consider the following situation.

I have a very important job interview, and I am looking for a shirt that goes well with my favorite pantsuit. Department stores' illumination is usually a mix of fluorescent lights and natural light coming from the windows (when there are any). These unusual lighting conditions can make 'objective' color identification quite hard. In order to decide whether the color of the shirt is indeed a color that can go with my pantsuit, I will probably have to move around the store and look at it both in full natural light *and* in full artificial light, maybe going back and forth multiple times.

Changing the context gives me information about what remains invariant as opposed to what changes, so that I can use that information to reach an accurate judgment about the shirt's 'objective' color. Even though, arguably, the fact that we can engage in this activity and reach the desired results at all depends on our capacity to experience constancy with respect to color, it seems premature to conclude that our success at eventually determining the color of the shirt depends our visual system representing that very color property. Experiencing a certain kind of invariance with respect to color plays a crucial role in my overall ability to reach an accurate color judgment, but jumping to the conclusion that the property attributed in judgment is itself represented in my perceptual experience, like the Absolute invariant view implies, is too quick and much less reasonable than it prima facie looks. Let me clarify this point further.

In order to explain how I manage to reach an accurate judgment about the color of the shirt, the tradition would say that my visual system analyzes the retinal image of the shirt in both contexts, separating the illumination 'factor' from those features that are intrinsic to the shirt's

surface on which its color depends. Isolating the 'real' color of the shirt, in this picture, takes multiple steps. First, the retinal image in context 1, call it R_1, has to be analyzed into a hypothetical illumination component (I_1) and a hypothetical 'real' color component (C_1). Second, this hypothesis ($R_1 = I_1 + C_1$) has to be 'tested' against the second analyzed retinal image in context 2 ($R_2 = I_2 + C_2$). Third, if the analysis of the new retinal image is consistent with the first one, that is, if the system 'establishes' that $R_1 - I_1 = R_2 - I_2$, then it 'concludes' that $C_1 = C_2$ = the 'real', context-independent color property possessed by the shirt.

On the other hand, RI would claim that retinal images aren't analyzed in this way, and that a representation of the context-independent color of the shirt by the visual system is not needed in order for such color property to be attributed to the shirt *in judgment*. The invariance experience with respect to color simply allows me to be visually aware of certain relations among multiple colored surfaces in the scene (presumably, the shirt isn't the *only* colored thing in your visual field at any given time) remaining unchanged in the two contexts. A judgment about the 'real' color of the shirt can be achieved later on, as a function of these invariant relations. Thus, differently from the traditional view, RI entails that the invariance experience is necessary but not sufficient for reaching a judgment about the 'real' color. On this view, experiencing invariance with respect to color only takes me part of the way: It 'informs' me that the global appearance change is consistent with the illumination changing and the 'real' colors staying the same. Finding out exactly what these 'real' colors are is a different task.

Even though I myself find the relational 'story' more intuitive and more in line with how I would pre-theoretically describe my experience of the situation above, the reader should not just take my word for it, or even rely on her own intuitions alone. Indeed, a new theory of color vision has been gaining popularity in both science and philosophy. Roughly, this theory claims that

representation of color relations, understood in particular as chromatic contrast properties (in the form of a 'holistic' iconic representation of edges separating colored surfaces) is more primitive than representation of monadic color properties (cfr. Davies, 2020). This view is supported by several studies involving patients diagnosed with cerebral achromatopsia, a condition that makes subjects visual experience completely achromatic. Despite their inability to see monadic colors, some achromatopsic patients are nonetheless apparently able to visually perceive local contrasts and discriminate among different appearances of edges (Kentridge et al., 2004a, 2004b). If this is right, then we might have a further reason to privilege a relational view of color constancy based on the very functioning of color vision.

On to the next everyday recognitional task: the (visual) re-identification of bodies, especially biological ones, in motion. As Green (2019) argues, because a moving biological body doesn't undergo rigid transformations, it is hard to account for our capacity to recognize it as the same body (in the sense of not changing its overall shape and size) in virtue of a specific representation of 'objective' monadic 3D shape. This consideration prompted Green to suggest that our capacity to recognize moving bodies across transformations depends on our visual system representing the underlying "compositional structure" of the body, which is a global, relational feature of physical bodies. [28] How would the tradition account for this invariance experience? I am going to argue that, due to its own theoretical assumptions, the traditional view in fact *cannot* account for the invariance experience of a body in motion.

It seems undeniable that, in order to perceive a body as moving through space – as opposed to undergoing other types of change, such as for instance changes in its intrinsic properties – we

[28] For more psychological studies supporting the claim that the visual system has a 'preference' for structure as opposed to monadic shape representation, see Guan & Firestone (2019) and Lowet, Firestone, & Scholl (2018).

must have a way to perceptually differentiate motion from these other kinds of transformations. Usually, that a body is moving rather than changing its intrinsic properties is something that's manifest in perceptual experience. There must be, therefore, a capacity (or set of capacities) of the visual system that allows for this difference to be experienced. The tradition implies that, in order for the patterns of change that characterize a body in motion to be experienced, certain specific properties of the body must be recovered and represented. In particular, the properties that would have to be recovered through a reverse inference-like process are shape, size, and location (Green calls them "geometrical constancies"). But this reverse inference-like process entails that any component of the proximal stimulus that's deemed to be perceiver- or context-dependent must be estimated and then discounted.

According to the traditional view, then, the invariance experience of a body in motion depends on perceiver- and context-dependent components of proximal stimuli being successfully discounted while the body's 'actual' 3D shape, size, and location are recovered. But if this is the case, then the tradition is committed to saying that the perceptual experience of bodies in motion does not involve contextual and perceiver-dependent information from proximal stimuli.

However, this seems clearly false. The experience of a moving object seems to depend on the object being experienced as changing its spatial location *with respect to the perceiver.* Motion perception requires representing the world ego-centrically. But representing the world egocentrically *requires* that contextual and perceiver-dependent information is taken into account. The tradition, therefore, faces a contradiction: on the one hand, it is committed to saying that contextual and perceiver-dependent information is irrelevant for the experience of a body in motion, though it seems clearly true that contextual and perceiver-dependent information is exactly what makes the experience of motion possible.

Moving outside of the visual domain, I have already discussed a relation-based account of loudness constancy with changing distance, as well as the case of timbre. Timbre in particular seems to be worth mentioning again in this context of recognitional tasks relying on constancy. Timbre recognition doesn't seem to fit the traditional account because the invariant feature that supposedly we recognize doesn't have a precisely identifiable physical correlate that's also entirely perceiver- and context-independent (analogous to, for instance, SSRs for colors).

If timbre isn't stably correlated with any specific perceiver- and context-independent property of any specific object, accounting for timbre constancy in traditional terms, that is, as precisely the capacity to recover a perceiver- and context-independent property of an object (presumably either the sound itself or the sound source) through stimulus analysis and reverse inference, seems impossible. On the other hand, an account of timbre recognition which focuses on the tracking by the auditory system of invariant patterns of (relevant) scene-wide relations seems more promising.

Consistently with this, the little empirical evidence collected so far favors the idea that timbre perception across different contexts has a lot in common with the relational model of color constancy. As we already saw in chapter 2 (section 2.2.2), timbre perception is affected by changes in the global soundscape analogously to how color perception is affected by changes in illumination (e.g. Stilp et al. 2010). The relational color constancy model considers illumination as a 'filter' affecting all colored surfaces in the same way, thus preserving the internal relations. Similarly, in the case of timbre, the auditory system standardly interprets changes (e.g. the addition of a new background sound/white noise, or a sound with certain specific spectral characteristics played before the presentation of the target timbre) as filters affecting the whole soundscape

uniformly, and 'adjusts' timbre perception according to the assumption that the internal acoustic relations won't be disrupted by the filter.

This bias towards preserving invariant acoustic relations within the soundscape (and the 'uniform filter' assumption whenever something changes) is also consistent with what we know about the auditory system more generally. Indeed, this system 'organizes' auditory stimuli by following roughly two basic 'rules': regularities can be ignored, irregularities must be 'over-emphasized' (e.g. Heald, Van Hedger, & Nusbaum, 2017; Kiefte & Kluender, 2008). In sum, the empirical study of timbre perception suggests two things: first, that timbre itself cannot be reduced to any perceiver- and context-independent property of a sound or a sound source; second, that timbre recognition depends heavily on whether the system can interpret the proximal auditory stimulus so that global acoustic internal relations are preserved. Far from being concerned with faithfully modelling the perceiver-independent acoustic relations among sound-waves, auditory systems structure the experienced soundscape in the way that provides subjects with most coherence and stability (cfr. Handel, 2006).

These were only three examples of recognitional tasks relying on invariance experiences (which in turn rely on perceptual constancy capacities), and for all three of them I suggested that we explain them in terms of tracking relational invariants instead of recovering absolute properties through reverse inference. Now, I will provide a little more evidence that perceptual systems are capable of tracking regularities in proximal stimulus change and, consequently, to represent relational invariants.

Ecological views of perception (e.g. Gibson & Crooks, 1938; Gibson, 1950, 1966, 1979; Lee & Lishman, 1975; Lee, 2011; Lee & Reddish, 1981) have been long committed to the idea that relations are both metaphysically and causally more basic than monadic properties when it

comes to accounting for how an animal embedded in an environment perceives. Ecological psychologists have devised lots of experiments aimed at showing how perceptually tracking (perceiver-involving) relations is what grounds and sometimes entirely replaces the attribution of (perceiver-independent) monadic properties to objects.

For example, quite convincing results have been obtained in the study of kinaesthetic touch and proprioception (e.g. Carello & Turvey, 2004; Jones, 2003). Pagano and Cabe (2003) showed that subjects can estimate the length of a hand-held rod by simply waving it around and relying on feedback from the muscles and tendons in their arm. The authors suggest that the proprioceptive and kinaesthetic systems are directly tracking a relational feature of the arm-rod system, namely its *moment of inertia*, i.e. an object's resistance to rotation along one dimension.

Another interesting case involves the so-called 'size-weight illusion' (Charpentier, 1891). The illusion consists in the fact that, given two objects of equal mass, generally the one with smaller diameter is judged to be heavier. Traditionally, the explanation was that the erroneous weight estimations depended on a subject 'over-compensating' in order to account for the smaller object's lower resistance to gravity. However, Amazeen and Turvey (1996) provided an alternative explanation. Instead of considering incorrect weight judgments in the illusion as the result of an erroneous mental computation, they showed how weight judgments of objects of equal size and weight varied according to the objects' different distribution of such weight. Since, when an object is held and moved around, weight distribution determines its moment of inertia in each direction of movement, Amazeen and Turvey concluded that perceived weight, which informed judgments in the size-weight illusion, most likely depends on perceived 'moveability' (Shockley, Carello, & Turvey, 2004), which in turn correlates with an object's inertia tensor (i.e. the object's resistance

to rotation in three dimensions at the wrist), that is, a complex relational feature of the object-perceiver system taken as a whole.

By studying the regularities emerging from a creature's perceptually guided interactions with her environment, ecological psychologists provided us with a set of mathematically expressible relational features of the creature-environment system – i.e. what Gibson (1979) called "invariants" – that a creature's perceptual system is directly sensitive to and relies on to guide action, with no need of further internal processing or conscious awareness that such relations are in fact being tracked.

In addition to guiding action, some of these relational features play also an important role in allowing the subject to have a stable and consistent conscious experience of herself and her surroundings. In particular, it has been shown how a stable visual experience of the world is intimately tied with our body slightly swaying back and forth when we stand and thanks to which we keep our balance when standing (e.g. Balasubramaniam et al., 2000; Lee & Lishman, 1975; Riley et al., 2011). This intimate connection is in part explained at the anatomical level. Indeed, the vestibular system is connected both with the neural structures controlling eye movement (thus being at the basis of the so-called vestibulo-ocular reflex, which is essential for clear and stable visual experience) and with the structures responsible for posture control (e.g. Khan & Chang, 2013). The connection is also evident at the psychological level, for instance in the phenomenon of height vertigo. Some people experience height vertigo, for example, when standing on top of a tall building and looking down. When there is a big distance between the observer and the surface observed, the very subtle bodily sway that is usually enough to keep us in an upright position is no longer enough to ensure stable visual experience. In order to restore the feedback loop between visual experience and postural sway, then, the vestibular system sends the signal to amplify the

postural sway, and that's why people report having the sensation of losing their balance and feeling as if they are about to fall off. Thus, both balance and stable perceptual experience seem to depend on each other and, more generally, they rely on the capacity of the perceptual system to keep track of certain stable relations obtaining between the perceiver and her surroundings.

All these considerations speak in favor of both the empirical and conceptual plausibility of the Relational Invariance view. However, more can be said regarding the more 'philosophical' component of the view. I will do this by discussing a potential phenomenological objection and by providing a reply to it.

4.3.1 The phenomenological objection

The objection goes like this: the phenomenal character of perceptual experiences like a wall as remaining of the same color throughout changes in illumination, or a table as remaining of the same shape when seen from different angles, seem to include the 'actual' shape that I am attributing to the table or the 'actual' color that I am attributing to the wall. It doesn't seem to be the case with relational invariants: what it is like to experience, say, a non-rigid body in motion as conserving its 3D shape does not include a representation of the body's compositional structure. Moreover, anecdotally, I can accurately attribute shapes and colors to objects on the basis of how they phenomenally appear to me. On the other hand, if asked to choose which 'absolute' compositional structure to attribute to a non-rigid body in motion purely on the basis of my experience, I wouldn't be as accurate. If what it's like to experience invariance in color is the same (i.e. the same *kind* of experience) as what it's like to experience invariance in compositional structure, how come that I can easily issue judgments on one but not on the other?

This is a quite common reaction in the context of discussions about the problem of Invariance, and it probably arises from the pre-theoretical tendency to think that we couldn't experience invariance in a property across contexts or viewpoints unless we were *also* at the same time aware of that very property. Put in another way, I take the thought here to be that the 'subjective' logically depends on the 'objective', so that something can appear to us in a certain way (i.e. the 'subjective') only insofar as we have access to the very thing that is appearing that way (i.e. the 'objective'). Because this idea is intuitive, and arguably the traditional view is usually considered more in line with pre-theoretical intuitions, it seems to work as a direct objection to RI. I will offer two replies; one quick, one more elaborate.

The first, quick reply is that, in cases like this, it is not clear whether one can determine through introspection alone what the actual phenomenal character of her experience is. When such familiar experiences are involved, it is difficult to disentangle how things 'actually' appear from how things 'should' appear in light of what we know, predict, or expect. Moreover, introspection in general is, according to some, far from reliable (e.g. Comte, 1830; Blackmore, 2002; Schwitzgebel, 2008, 2012). Also, consider that the Scottish psychologist Robert Thouless first noticed how subjects, when asked to report how an object appeared from their current perspective and in the current context, systematically reported it to look much closer to how it 'objectively' was than to how it should have appeared when projected on the retina. Thouless calls this phenomenon the "phenomenal regression to the 'real' object" (1931, 344). However, there is evidence in developmental psychology suggesting that such a regression, rather than phenomenal, is in fact cognitively mediated, that is, subjects 'read' their own experience through the lenses of

what they know or believe about the situation. Although we can be sincere in claiming that this is how things look to us, this doesn't necessarily entail that they do, in fact, look that way.[29]

The difficulty in 'breaking down' our own experiences, and especially in introspectively distinguishing the 'subjective' elements from the 'objective' ones, was the main reason why some philosophers interested in the problem opted for different solutions altogether. On the one hand, we find 'cognitivist' views such as, for example, Bertrand Russell's (1912), who argued that objects in perception appear only in their context- and perceiver-dependent aspect, while the 'objective' aspect is added by knowledge/belief. Earlier on, Hering (1878) had proposed a similar view, but based on memory rather than belief: whenever presented with a certain appearance/set of appearances, we retrieve from memory the actual properties that generated those same appearances before, and we 'interpret' the current appearances in light of the memory.

On the other hand, naïve realists such as Smith (2002) claim that we perceptually experience the 'objective' properties of objects while what is subjective and perspectival is merely causally responsible for the experience of objects 'as they really are'. This follows from the naïve

---

[29] Another, more 'conciliatory' line of reply is worth mentioning; one which neither implies that introspection is utterly unreliable. This consists in generalizing a point made by Will Davies (2016) regarding color constancy and color phenomenology. First, however, a bit of context is needed. Recall that the 'phenomenological objection' to RI that we are discussing right now goes like this: When displaying perceptual constancy, the objective, perceiver-independent properties of objects are present in the phenomenal character of my experience. This idea, in turn, seems to rely on a stronger claim regarding the nature of perceptual constancy; namely, that perceptual constancy consists in the invariance of an object's (or one of its properties') *appearance* across contextual changes. For example, a red cup in different illuminations, even though some changes in the experience are noticed, nonetheless keeps *phenomenally appearing* that very same shade of red to the observer. This claim in the particular case of color constancy is what Davies calls 'Phenomenal invariantism': "color constancy consists in some invariance in the phenomenology of color experience, that is, in the qualitative dimensions of conscious *color appearance*" (2016, 541; emphasis in the original text). Davies favors a more 'nuanced' version of invariantism, which he calls 'Complex invariantism', based on the distinction of two 'aspects' of color appearance called 'material' color appearance and 'lighting' color appearance (cfr. Mausfeld, 2003): "color constancy consists in the approximate invariance of material color appearance, while allowing for significant variation in lighting color appearance." (Davies, 2016, 552). Davies, then, uses this view to claim that "the phenomenological basis for color constancy should be assessed on a case-by-case basis, with appearance invariance […] playing more or less of an explanatory role, depending on context" (2016, 555). Now, even though Davies doesn't explicitly make this last step, I take the lesson here to be that the appeal to color phenomenology is in general not a reliable way to support claims about the nature of color constancy. This lesson generalizes to other perceptual constancies, and can be used to resist at least in part the phenomenological objection.

realist commitment to the idea that perceptual experiences (and their phenomenal character) are constituted by the actual objects in the world and their actual properties, and this entails that the perceiver- and context-dependent elements cannot themselves be phenomenally experienced. Both the cognitivist and the naïve realist positions have been challenged by advocates of the traditional view on precisely phenomenological grounds, by observing that there is clearly a sense in which both 'how the object appears from this particular viewpoint/in this particular context' and 'how the object appears to *be*' (i.e. the properties that we would attribute to it in judgment) are part of our ordinary perceptual experiences (e.g. Noë, 2012).

I will now argue that the phenomenological objection can be resisted precisely on phenomenological grounds, and that, in fact, a closer look at what it's like to have invariance experiences ultimately favors RI over the tradition. For my argument, I will draw on one of the richest and most thoughtful phenomenological descriptions of perceptual experience ever attempted, which we owe to Merleau-Ponty (1964, 2013; Matherne, 2017).

Merleau-Ponty thinks that perceptual experience – and in fact, all conscious experience – is characterized by what he calls the "paradox of immanence and transcendence" (Merleau-Ponty, 1964, 16).[30] This means that, through perception, we are at once occupying our own, immanent and contingent point of view on the world and also capable of 'transcending' it to experience objects in their objective, mind-independent nature. In particular, Merleau-Ponty thinks that his way of describing the phenomenology of perceptual experience will illuminate both its 'subjective' and 'objective' components and, re-elaborating a traditionally Kantian idea, show how, far from constituting a paradox, the simultaneous presence of both components is the very condition of

[30] The fact that Merleau-Ponty talks in terms of a 'paradox' shouldn't be taken too seriously: all that he means, at least on my reading, is that it *could* appear like a paradox until we provide a proper phenomenological description of the situation.

possibility of experience. Recall that, in order to reply to the phenomenological objection on phenomenological grounds, the advocate of RI must be able to show how the phenomenal character of *all* our invariance experiences necessarily includes an invariant relational component. I believe that Merleau-Ponty has provided us with such a component, namely what he calls "style".

An object's style is its unified manner of being, that is, a sort of global 'aura' that uniquely determines that individual object. Merleau-Ponty explains that "the perceived thing is not an ideal unity in the possession of the intellect, like a geometrical notion, for example; it is rather a totality open to a horizon of an indefinite number of perspectival views which *blend with one another* according to a given style, which defines the object in question" (Merleau-Ponty, 1964, 16; emphasis mine). This isn't too far away from our intuitive concept. Think about a painter or a musician: often, upon seeing multiple works from an artist, we start grasping something that is shared by all the works, though not quite in the same way.

The style of an artist is precisely that quite ineffable and irreducible feature that all of her works share, in various degrees and by means of various specific material features. Works of art have styles, which are usually expressions of the style of their creator, and we have the capacity to recognize styles without necessarily associating them to any particular property of any specific work. People themselves, in virtue of being unique individuals, have styles, too. According to Merleau-Ponty, we recognize a person by means of recognizing her particular style before we perceive any specific detail about them (e.g. the clothes they are wearing, the expression on their face, their haircut, etc.).

Therefore, on this view, the experience of an object' style precedes and grounds every other perceptual experience of the object, including the experience of any of its 'absolute' properties, similar to how a musical 'theme' can unify and ground the identity of a musical composition (e.g.

a symphony) though the theme gets 'developed' and revisited in many different fashions throughout the work. It can be that, in a musical piece, there is never a perfect repetition of the same succession of notes, with the same tempo, or played by the same instrument; yet, the style of the piece is clearly recognizable throughout. Styles, for Merleau-Ponty, are non-repeatable and non-reducible to specific patterns or 'laws', thus being different from things like 'gestalten'. Nevertheless, they are systematically and reliably recognized by perceivers throughout contexts and over time, that is, they ground and sustain invariance experiences. Connecting back to the phenomenological objection, here is how Merleau-Ponty's notion of style provides a reply to it.

Merleau-Ponty's notion of style accounts for *all* invariance experiences, including those easily captured by the traditional approach. Indeed, according to Merleau-Ponty, we can recognize specific properties of objects like shapes or colors only *after* (or, more precisely, *in virtue of* the fact that) we have recognized the style. This is because, on this view, absolute properties are not perceived separately and subsequently bound together to form individual objects; rather, they are "abstract moments of the constancy of things" (Merleau-Ponty, 2013, 326). The invariance experience of the object (or even a whole scene!) as a unity characterized by a certain style precedes and influences the experience of any single property we might attribute to that object. Merleau-Ponty uses the example of having the invariance experience of a blue woolen rug under different illuminations: "A color is never simply a color, but rather the color of a certain object, and the blue of a rug would not be the same blue if it were not a wooly blue. […] The colors of the visual field […] form an ordered system" (Merleau-Ponty, 2013, 326).

Samantha Matherne (2017) shares the same view, and I largely borrow this account of Merleau-Ponty's position from her:

> Each property […] is determined by the object's style insofar as it is an expression of it. Moreover, he thinks that this style is something that unifies that property with

> all the other properties in an object and that this further determines the property. […] we cannot grasp the color of the carpet without grasping the other properties of the carpet, these other properties are, in turn, unified together in virtue of the carpet's […] style.
>
> […] if you have two red objects, say a red carpet and a red dress, the redness of each will not be the same because each red will be determined by the respective object's style and the way in which this style binds the redness to the object's other properties.
>
> […] constant properties do not float free from, but rather depend on, an object and its style. (Matherne, 2017, 719-720)

Overall, I find Merleau-Ponty's (and Matherne's) account of the phenomenology of invariance rooted in the notion of style very plausible. It fits nicely with the Relational Invariant view, and provides an explanation of both how it can sometimes still appear like we directly experience the absolute perceiver-independent properties of objects: the recognition of a style opens the possibility for us to reflect on the object further and isolate its absolute properties. These properties, however, enter our conscious experience as a result of intellectual analysis rather than as originally part of the perceptual synthesis. This idea allows a defender of RI to resist the phenomenological objection on phenomenological grounds.

4.4 Intermediate summary

We experience a stable world; stable enough for us to (i) judge accurately which properties objects have and (ii) act successfully, among other things. We want to explain this stability, and the capacity of perceptual constancy seems to be at the core of the issue. Indeed, the consensus is that it is in virtue of this capacity if perceptual systems that perceptual inputs are stabilized and that this stability can also, though not necessarily, be experienced. But once we have clarified the

‘job’ of constancy, the question becomes: What does constancy consist in? What is a perceptual system doing whenever we say that it is ‘displaying’ or ‘achieving’ constancy? The tradition answers: “The perceptual system solves the underdetermination problem and recovers accurate representations of the perceiver-independent properties of distal objects.” To display, for instance, shape constancy, is to recover a representation of the actual shape of an object by discounting whatever is taken to be contextual or perceiver-dependent. On the other hand, the answer given by the Relational invariance view is: “[when a perceptual system displays constancy], it tracks invariant relational features of the scene thanks to regularities in how proximal stimuli change over time”. No specific perceiver-independent property needs to be represented in order to reach and maintain perceptual stability.

From how the Relational Invariance view defines perceptual constancy, it follows that there are in fact many more perceptual constancies than the Traditional view assumed, and that these capacities may involve more than one sense modality at once. After all, objects have only a limited number of modality-specific perceivable monadic properties, but they can stand in an infinite number of relations with each other and with the perceiver. These relations, then, can ground perceptual invariance even without the perceiver consciously experiencing such invariance, as long as it can be relied upon to guide motor action. This is something that follows from, on the one hand, the observation that it is hard to deny that constancy is involved in perceptually guided motor action and, on the other hand, that perceptual guidance of motor action might not be conscious, as suggested by the separation of dorsal and ventral stream in the visual system.

In sum, the Relational invariance view re-thinks quite radically what perceptual constancy is constancy ‘of’. But if one is open to re-considering what this capacity looks like, especially in light of its being arguably necessary for perceptually guided motor action, everything else falls

into place as a consequence. I have already shown that we have philosophical reasons to change framework, which derive from what we think the *role* of a capacity like constancy should be. This, I think, is sufficient to establish that the new framework is worth exploring further.

More generally, the way one understands psychological capacities is always a matter of what kind of theoretical commitments one takes up: at this level, there is no 'theory-neutral' description of how things are. I think that the case of perceptual constancy is emblematic in this sense: a psychological capacity and a set of mechanisms is postulated in order to explain a certain set of phenomena that are part of everyday perceptual experience, but then the study of this capacity and the class of cases it's supposed to explain are 'colored' by theoretical assumptions regarding the nature of perception more broadly, experimental methodology, and more. As a result, the theoretical framework in which the capacity is defined and studied controls and limits our understanding of the capacity, allowing for explanatory 'blind spots'. For the traditional view, the role of constancy in perceptually guided motor action lies exactly in this kind of 'blind spot', or so I argued. But once this 'diagnosis' has been put forward, then it is up to theorists to decide whether to stick to their assumptions and keep understanding constancy the traditional way, or to be open to exploring something different. My own view on the matter is clear enough at this point. To say it with John Dewey (1896): "It is always wise to beware of that false simplicity which is reached by leaving out of account a large part of the problem". Now it is time to move to objectivity, and to finally put all the pieces together.

4.5 A 'coda' on objectivity

Consider the following definition of what it means to perceive objectively, or to 'objectify': "to objectify is to perceive more than is provided by the stimulus, and thus to definitely move beyond forms of responding that are directly guided by mere sensory stimulation" (Hutto & Myin, 2013, 119). So far, we have seen constancy being defined as the capacity to 'go beyond' what's directly available through proximal stimulation, whether this means de-composing the stimulus to extract a representation of a perceiver-independent property or exploiting emerging regularities in how the stimulus changes over time. Now, this 'going beyond' idea is used to get a grip on what perceiving objectively means, too. This raises the questions: Is objectivity the same as constancy? Are the two notions interchangeable? The answer is no, but these questions have a point nonetheless. Indeed, the main difference between the traditional framework and the framework resulting from the Relational invariance view of constancy is how tight they assume the relationship between constancy and objectivity to be. Even though neither view claims that they are interchangeable, there is a clear sense in which the traditional framework considers them more directly connected than the Relational Invariance framework does. The goal of this section is to explore these issues and ultimately clarify what perceiving objectively amounts to if one accepts the Relational Invariance view of perceptual constancy.

Once again, we owe the formulation of the traditional view of objectivity to Burge. This view of objectivity understands this notion as 'neutrality', as the result of a perceiver 'stepping out' of the environment in order to get a better sense of how the world is independently of how she experiences it 'from within'. This is the very same idea behind the traditional account of constancy: perceptual systems must find a way to compensate for partial, ambiguous proximal stimuli in order to reach a perspective-independent representation of the actual properties

possessed by objects. Getting rid of and compensating for 'idiosyncrasies' is the core of Burge's account of what it is to 'objectify' in perception:

> Objectification, then, is the formation of a representational state that represents the physical environment, beyond the individual's local, idiosyncratic, or subjective features. (2010, 397)
>
> Specification of mind-independent and constitutively non-perspectival physical entities is separated out from the individual's sensory registration—the functioning state that encodes proximal sensory information. Perceptual states are products of such systematic separation and privileging processes. (398)
>
> Objectification here hinges on distinguishing and contrasting, in the operations of the system, what concerns the individual's receptors and what concerns a receptor-independent reality—and doing so in an attribute-specific way. (398)
>
> […] objectification separates local, idiosyncratic registrations from representations of individual-independent, occasion-independent, mind-independent, perspective-independent reality, beyond the individual. (399)

Burge acknowledges also that this 'stepping out' of perspective to represent the real way in which the world is, abstracting from all idiosyncrasies, cannot be as 'perfect' in perception as it is, for example, in theoretical science. Nevertheless, we should accept the idea that perception, when perceptual systems display constancy, is objective, or at least as objective as it can be. Even though every objective representation is necessarily causally related to the egocentric features of the scene and, thus, to proximal stimuli, this isn't enough to jeopardize perception's capacity to be objective:

> The objectivity of perception should not be denigrated because it is not the context-independent objectivity of theoretical science. Perception is inevitably from the perceiver's perspective. It is constitutively marked by egocentric representational frameworks. (2010, 401)

Objectivity, in the traditional framework, means neutrality; it means independence from any influence of contingent factors like context, idiosyncrasies of individual perceivers (and their perceptual systems), perspective, etc. However, if one rejects the claim that reaching a

representation of the one and only way in which the world *really* is by overcoming the ambiguity of proximal stimuli is what perceiving objectively requires, then the association of objectivity with neutrality becomes less compelling. There are other ways to understand the process that takes from ambiguous and fluctuating proximal sensory stimuli to perceptual representations of the world which can successfully guide a creature's behavior and, as a consequence, there are other ways to understand objective perception.

In particular, there is the way in which the Relational Invariance view of perceptual constancy understands the relationship between proximal sensory stimuli and action-guiding (and judgment-grounding) percepts. The story told by the Relational Invariance view of constancy has implications for how we understand objectivity, and the goal of this section is to articulate how exactly it is so. The first step is to replace the notion of neutrality with another term that better captures the 'spirit' of objectivity in the new picture. I suggest that we use the notion of *authenticity*: the term 'objective' can be, instead of synonymous with 'neutral', synonymous with 'authentic'.

Think about, for example, the way we discuss whether some journalist is 'objective' in the way they talk about a particular issue. In this context, being objective might mean being neutral, that is, not taking any 'side' and simply 'sticking to the facts'. However, there is also a sense in which 'objective' can mean that the journalist is not refraining from interpretations altogether, but rather is trying to figure out which side is *right,* who is telling the *truth.* This second way to understand the journalist's objectivity seems closer to the idea of authenticity, or what we might call 'intellectual honesty'. This association of objectivity with authenticity can be employed to the context of perception, too: there is a sense, which the tradition would probably reject, in which perceiving objectively means perceiving authentically, that is, in the 'right' way for a creature

having to survive in a particular environment. This will mean that, likely, there are multiple 'objective' ways to perceive the world, depending on what the 'truth' for a particular creature in a particular environment is. A valuable resource to start articulating this different conception of objectivity in perception is the work of Maurice Merleau-Ponty (1964, 2013).

In complete contrast with Burge, Merleau-Ponty thinks that perspectivalness, egocentricity, context-dependence etc. are features, rather than bugs, when it comes to perceiving objectively. After all, if perceiving objectively is what perception is *for*, then how can something that is part of the intrinsic nature of perception be considered an obstacle? He writes: "It is not accidental for the object to be given to me in a 'deformed' way, from the point of view which I occupy. That is the price of its being 'real'." (Merleau-Ponty, 1964, 15-16). Perhaps, and this is Merleau-Ponty's idea, we have largely misunderstood what objectivity in perception can be in the first place, thus putting implausible 'demands' on it. The result of this misunderstanding is seeing it as somehow 'imperfect', like Burge seems to do. Perceptual objectivity, according to Merleau-Ponty, is of a fundamentally different *kind* from the objectivity achieved in science, or through abstract thought. One fundamental difference is that, in perception, objectivity must not be "confused with what is measurable" (1964, 24). This is a clear criticism of the view in psychology and philosophy of perception which conceives of perceptual systems as measuring instruments (I discussed and criticized this view in chapter 3):

> Is one truly objective with respect to man when he thinks he can take him as an object which can be explained at the intersection of processes and causalities? Is it not more objective to attempt to constitute a new science of human life based on the description of typical behaviors? Is it objective to apply tests to man which deal only with abstract aptitudes, or to attempt to grasp the situation of man as he is present in the world and to other by means of still more tests? (1964, 25)

In what sense is perception objective, then? Merleau-Ponty's answer sounds, at least prima facie, paradoxical: the objectivity of perception necessarily depends on its intrinsic *indeterminacy*.

But how can indeterminacy afford objectivity? Intuitively, when we say that something is perceived objectively, we seem to think that that thing is perceived as determinate, 'final', as opposed to what's perceived subjectively, that is, in a way that's affected by our own flaws and limitations. The commonsense view, therefore, seems to take indeterminacy to be the 'enemy' of objectivity: what's indeterminate is usually understood as confused, temporary, waiting to be made determinate. However, according to Merleau-Ponty, there is no objective perception without indeterminacy, because indeterminacy simply is part of the very nature of perception.

Consider this example. A photograph of a landscape 'reduces' it to a flat, two-dimensional surface, where everything is determinate because nothing is hidden: the picture of a tree does not have a back side, while a real tree does. In a picture, by just looking at it we see all there is to be seen. In the same way, when we judge, or we represent the perceiver-independent properties of objects, we can think of ourselves as 'taking a picture' of what's in front of us, thereby making the scene determinate. On the other hand, the real world is made up of three-dimensional objects, and for this reason it is not all determinate: at any given time there is something we are not seeing but that, at the same time, necessarily influences what we do see. These are the 'hidden' aspects of objects, i.e. those that are not in direct contact with our sense organs at each instant. Perceiving the world is different from looking at a picture of it because the former, and only the former, has something necessarily indeterminate to it.

Perception of the world just is *structured* that way: there is always co-presence of figure, i.e. the specific object we are perceiving, the 'focus' of our attention, and background, i.e. the context in which the 'main' object is presented, including other objects we're not currently attending to, and the perspectival aspects that are currently invisible from our point of view: "Objects form a system in which one object cannot appear without concealing others. More

precisely, the inner horizon of an object cannot become an object without the surrounding objects becoming a horizon, and so vision is a two-sided act." (Merleau-Ponty, 2013, 70). The figure-background structure is necessary for objectivity: standing against a background is part of the definition of an *object* (of perception). Something can be figure only when experienced against a background, and these are the conditions for objective perception, that is, perception of objects *as objects*, entities that exist independently of us.

Merleau-Ponty's view, thus, allows us to flip the idea of objectivity we got from Burge and the tradition upside down: in order to perceive objects, i.e. to perceive objectively, I don't have to 'step outside' and form a perspective-independent 'model' of those objects. Rather, objectivity is achieved by 'stepping further in', by inhabiting the world, engaging with it, moving around to disclose more of it and make regularities emerge:

> To see an object is to come to inhabit it and to thereby grasp all things according to the sides these other things turn toward this object. And yet, to the extent that I also see those things, they remain places open to my gaze and, being virtually situated in them, I already perceive the central object of my present vision from different angles. […] Thus, I can see one object only insofar as objects form a system or a world, and insofar as each of them arranges the others around itself like spectators of its hidden aspects and as the guarantee of their permanence. (2013, 71)

An object, according to Merleau-Ponty, is not an object unless it is experienced from a perspective; if we are perceiving a real object in the world out there, then this is hiding something from us: "the thing and the world only exist as lived by me, or as lived by subjects like me" (2013, 349). And to 'live the world' means to use our body and explore it. To perceive objectively means to be sensitive to the indeterminacy and 'openness' of objects: "it is essential for the thing and for the world to be presented as 'open', to send us beyond their determinate manifestations, and to promise us always 'something more to see'." (348)

With this account of objectivity in mind, let's go back to constancy and see how the Relational Invariance view fits with what Merleau-Ponty argued. The general idea is that, in the relational framework, perceiving an objective world doesn't mean perceiving the actual properties of specific objects and, in virtue of that, experiencing them as mind-independent. Rather, perceiving objectively means experiencing both objects and oneself as part of a *system*, immersed in a web of relations with the environment. These relations need not be experienced all the time (I have discussed this at length above) but perceiving objectively necessarily requires consciousness and, to an extent, self-consciousness. The perceived world is an objective world insofar as (and because) it is experienced *by me,*[31] that is, if I can consciously be aware both of myself as the center of my experience, and of the world as what makes that very experience possible and provides constraints to my movements and activities. The 'self' necessary for objective perception is mostly a *bodily* self, that is, the coherent and unified system of bodily habits and dispositions to interact with the environment through movement. The different sense modalities are all organized in a way that aims at coherence and unity, too, so that objective perceptual experience is 'global' and relies on the *structures* connecting the whole scene and the perceiver, rather than on local components such as single objects or single properties. In the Relational invariance framework, objectivity is achieved when the perceiver-environment relations – whose invariance is tracked by perceptual systems' constancy mechanisms – are consciously perceived, created, and controlled.

Therefore, in the Relational Invariance framework, constancy is necessary, but not sufficient for objectivity: constancy displays must be accompanied by a conscious awareness of

[31] It should be noted that neither Merleau-Ponty nor I are claiming that each perceiver is 'trapped' in her own experiential world which makes sense only to her and for her. *Intersubjectivity* is part of Merleau-Ponty's notion of objectivity because of the emphasis on relations: relations with other subjects are just as important as relations with objects.

them, though it isn't part of the very nature of constancy that its displays are consciously experienced. In the traditional framework, on the other hand, conscious awareness was already among the necessary requirements for constancy, and for this reason traditional constancy and objectivity are more immediately connected. According to the traditional view, indeed, displaying constancy at the level of perceptual systems (i.e. recovering perceiver-independent properties by solving the 'standard' inverse problem) is all it takes to sustain objective perception at the personal level. As Burge himself claims: "Perceptual constancies are capacities for objectification" (2010, 399).

One last question that should be answered is: What do we gain from endorsing the Relational Invariance framework (involving both the Relational Invariance view of constancy and the view of objectivity we largely owe to Merleau-Ponty)? Why do we need an alternative to the tradition in the first place? To answer this question, it is useful to look more closely at the *Perceptual Invariance* thesis*:*

Perceptual Invariance: Perceptual constancies are the capacities of perceptual systems responsible for the kind of perceived invariance *necessary* for stable conscious perceptual experience, accurate perceptual judgments, and successful perceptually guided motor actions.

The main reason why we need a new framework is the following: If one accepts *Perceptual Invariance* and some of its implications, then the traditional view of constancy is untenable. Moreover, if we want to preserve the connection between constancy and objectivity Burge insists on, a new view of objectivity is required, too. *Perceptual Invariance* is meant to be intuitively true. We need perceptual constancy in order to form beliefs and perform actions on objects on the basis of what we perceive.

Now, the traditional view holds that perceptual constancy is the capacity to recover perceiver-independent properties by discounting contextual and perceiver-dependent changes in

the proximal stimulus. If we accept *Perceptual Invariance*, it follows that recovering these perceiver-independent properties is *necessary* for successful motor action. However, this last claim is false, at least in the visual case: patient D.F. is a counterexample. Because she can perform visually guided motor actions, she can also display constancy (once again, this follows from *Perceptual Invariance*), and this is enough to conclude that the traditional view is in need of revision. The traditional view cannot account for constancy in action: it would either claim that D.F. doesn't really display constancy, thus rejecting the claim that constancy is necessarily enabling perceptually guided motor actions, or take the 'two capacities' route and endorse the relational invariance account only for perceptually guided action.

However, I have argued that we would be better off overall by rejecting the traditional view and endorse the Relational Invariance view to account for all types of constancy in all circumstances. The Relational Invariance view explains the role of constancy in action and, moreover, has the resources to accommodate most of the evidence on which the traditional view relied. And once the need for a new view of constancy is established, the next step is easy: objectivity can no longer be understood consistently with the traditional view of constancy the way Burge does. As a consequence, we need a different view of objectivity compatible with the Relational Invariance view of constancy, and it turns out that Merleau-Ponty had already done that for me, for the most part.

4.6 Conclusion

In this chapter, I presented the Relational Invariance view of constancy. On the one hand, this view explains how constancy can be responsible for the kind of perceptual invariance needed

for successful perceptually guided motor action. On the other hand, it is compatible with most of the ideas that motivated the traditional view. The examples discussed in section 4.3. meant to show that there is a vast range of invariance experiences in various sense modalities, and that they all share one common purpose: they allow us to recognize and re-identify objects across contexts and in non-ideal circumstances. Moreover, looking more closely at what could enable and explain successful recognition at the level of perceptual systems, tracking of relational invariants seems a more promising hypothesis than the recovery of absolute perceiver-independent properties of specific objects.

For example, I argued that the traditional approach has trouble accounting for the recognition of bodies in motion, which relies on an invariance experience. In a non-visual sense modality, I gave the example of timbre recognition, which also relies on an invariance experience, and does not seem to fit an explanation in terms of absolute perceiver-independent properties. I also briefly presented some results obtained by ecological psychology, which point in the direction of relational invariants being at the core not only of many perception-based recognitional tasks, but of perception itself. Then, I presented a phenomenological objection that seems to prima facie favor AP over RI. However, through Merleau-Ponty's notion of style, I replied to the objection and argued that we should prefer RI precisely because it offers a better account of the phenomenal character of invariance experiences.

Lastly, I connected RI to an account of what it means for perception to be objective. On the one hand, this account, rooted in Merleau-Ponty's theory of perception, still preserves the connection between displaying constancy and perceiving objectively that is at the core of Burge's account but, on the other hand, changes our understanding of objectivity consistently with the way I argued we should change our understanding of constancy. Overall, the Relational Invariance

framework, with its perhaps counter-intuitive accounts of both constancy and objectivity, invites us to re-think how living creatures with perceptual systems relate to their environment, and how they accomplish the incredible variety of tasks requiring stable perceptual inputs for motor action and an experience of the world as mind-independent.

Bibliography

Adams, Z. and Maher, C. (2017), "Constancy mechanisms and the normativity of perception", in Z. Adams and J. Browning (eds.) *Giving a Damn: Essays in Dialogue with John Haugeland,* MIT press, pp. 243 – 268.

Akins, K. (1996), "Of sensory systems and the 'aboutness' of mental states", *Journal of Philosophy*, 93, pp. 337–372.

Amano, K., Foster, D. H., & Nascimento, S. M. C. (2005). Minimalist surface-colour matching. *Perception*, *34*(8), 1009–1013. https://doi.org/10.1068/p5185

Amano, K., Foster, D. H., & Nascimento, S. M. C. (2006). Color constancy in natural scenes with and without an explicit illuminant cue. *Visual Neuroscience*, *23*, 351–356.

Amazeen, E. L., & Turvey, M. T. (1996). Weight perception and the haptic size–weight illusion are functions of the inertia tensor. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 22, pp. 213–232. https://doi.org/10.1037/0096-1523.22.1.213

Balasubramaniam, R., Riley, M. A., & Turvey, M. T. (2000). Specificity of postural sway to the demands of a precision task. *Gait & Posture*, *11*(1), 12–24. https://doi.org/https://doi.org/10.1016/S0966-6362(99)00051-X

Barwich, A.-S. (2019). A Critique of Olfactory Objects . *Frontiers in Psychology* , Vol. 10, p. 1337. Retrieved from https://www.frontiersin.org/article/10.3389/fpsyg.2019.01337

Barwich, A.-S. (2020). *Smellosophy - What the nose tells the mind*. Cambridge, MA: Harvard University Press.

Blackmore, S. (2002). There Is No Stream of Consciousness. *Journal of Consciousness Studies*, *9*, 17–28.

Boundy-Singer, Z. M., Saal, H. P., & Bensmaia, S. J. (2017). Speed invariance of tactile texture perception. *Journal of Neurophysiology*, *118*(4), 2371–2377. https://doi.org/10.1152/jn.00161.2017

Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America A*, *15*(2), 307–325. https://doi.org/10.1364/JOSAA.15.000307

Brown, R. (2003). Backgrounds and illuminants: the yin and yang of color constancy. In R. Mausfeld & D. Heyer (Eds.), *Colour Perception – Mind and the physical world* (pp. 247–272). Oxford: Oxford University Press (OUP).

Burge, T. (2005). Disjunctivism and Perceptual Psychology. *Philosophical Topics*, *33*(1), 1–78.

Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press (OUP).

Carello, C., & Turvey, M. T. (2004). Physics and Psychology of the Muscle Sense. In *Psychological Science* (Vol. 13).

Cartwright, N. (1983). *How the Laws of Physics Lie*. https://doi.org/10.1093/0198247044.001.0001

Charpentier, A. (1891). Analyse experimentale de quelques elements de la sensation de poids. *Archives Physiologiques Normales and Pathologiques*, *18*, 79–87.

Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT press.

Chirimuuta, M. (2016). Vision, Perspectivism, and Haptic Realism. *Philosophy of Science*, *83*(5), 746–756. https://doi.org/10.1086/687860

Chirimuuta, Mazviita. (2015). *Outside Color: Perceptual Science and the Puzzle of Color in Perception*. Cambridge, MA: MIT press.

Chirimuuta, Mazviita. (2017). Perceptual Pragmatism and the Naturalized Ontology of Color. *Topics in Cognitive Science*, *9*(1), 151–171. https://doi.org/10.1111/tops.12222

Chirimuuta, Mazviita. (2018). Vision. In M. Sprevak & M. Colombo (Eds.), *The Routledge Handbook of the Computational Mind*. London: Routledge.

Chirimuuta, Mazviita. (2019). Charting the Heraclitean Brain: Perspectivism and Simplification in Models of the Motor Cortex. *Understanding Perspectivism: Scientific Challenges and Methodological Prospects*, 141–159.

Davies, W. (2020). Colour Relations in Form. *Philosophy and Phenomenological Research*, *00*, 1–21. https://doi.org/10.1111/phpr.12679

de Almeida, V. M. N., Fiadeiro, P. T., & Nascimento, S. M. C. (2010). Effect of scene dimensionality on colour constancy with real three-dimensional scenes and objects. *Perception*, *39*(6), 770–779. https://doi.org/10.1068/p6485

Deubel, H., Koch, C., & Bridgeman, B. (2010). Landmarks facilitate visual space constancy across saccades and during fixation. *Vision Research*, Vol. 50, pp. 249–259. https://doi.org/10.1016/j.visres.2009.09.020

Dewey, J. (1896). The Reflex Arc Conceot in Psychology. *The Psychological Review*.

Foster, D. H. (2003). Does colour constancy exist? *Trends in Cognitive Sciences*, *7*(10), 439–443. https://doi.org/10.1016/j.tics.2003.08.002

Foster, D. H., & Nascimento, S. M. C. (1994). Relational Colour Constancy from Invariant Cone-Excitation Ratios. *Proceedings of the Royal Society B: Biological Sciences*, *257*(1349), 115–121.

Gibson, J J, & Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *The American Journal of Psychology*, *51*, 453–471. https://doi.org/10.2307/1416145

Gibson, James J. (1950). *The Perception of the Visual World*. Boston, MA: Houghton Mifflin.

Gibson, James J. (1966). *The Senses considered ad Perceptual Systems*. Boston, MA: Houghton Mifflin.

Gibson, James J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Giere, R. N. (2006). *Scientific Perspectivism*. Retrieved from http://ebookcentral.proquest.com/lib/pitt-ebooks/detail.action?docID=471866

Gilchrist, A. (2012). Objective and subjective sides of perception. In *Visual experience: Sensation, cognition, and constancy.* (pp. 105–121). https://doi.org/10.1093/acprof:oso/9780199597277.003.0006

Godfrey-Smith, P. (2016). Mind, Matter, and Metabolism. *Journal of Philosophy*, *113*, 481–506. https://doi.org/10.5840/jphil20161131034

Goodale, M. A., & Milner, D. A. (2013). *Sight Unseen. An exploration of conscious and unconscious vision* (2nd ed.). Oxford: Oxford University Press (OUP).

Granrud, C. E. (2009). Development of size constancy in children: A test of the metacognitive theory. *Attention, Perception, & Psychophysics*, *71*(3), 644–654. https://doi.org/10.3758/APP.71.3.644

Granrud, C. E. (2012). Strategy use in size judgments. In G. Hatfield & S. Allred (Eds.), *Visual Experience – Sensation, Cognition, and Constancy* (pp. 13–34). Oxford: Oxford University Press (OUP).

Granrud, C. E., & Shmechel, T. N. (2006). Development of size constancy in children: A test of the proximal mode sensitivity hypothesis. *Perception and Psychophysics*, *68*, 1372–1381.

Guan, C., & Firestone, C. (2019). Seeing What's Possible: Disconnected Visual Parts Are Confused for Their Potential Wholes. *Journal of Experimental Psychology: General*, *1*(999), 1–9. https://doi.org/10.1037/xge0000658

Handel, S. (2006). *Perceptual Coherence: Hearing and Seeing*. Retrieved from https://books.google.com/books?id=Y_WekgEACAAJ

Hatfield, G. (2009). On Perceptual Constancy. In G. Hatfield (Ed.), *Perception and Cognition: Essays in the Philosophy of Psychology* (pp. 178–211). Oxford: Oxford University Press (OUP).

Haugeland, J. (1996). Body and World: A Review of What Computers Still Can't Do: A Critique of Artificial Reason (Hubert L. Dreyfus). *Artificial Intelligence*, *80*, 119–128.

Heald, S. L. M., Van Hedger, S. C., & Nusbaum, H. C. (2017). Perceptual plasticity for auditory object recognition. *Frontiers in Psychology*, *8*(MAY). https://doi.org/10.3389/fpsyg.2017.00781

Helmholtz, H. von. (1867). *Handbuch der physiologischen Optik*. Leipzig: Leopold Voss.

Hochberg, J. (1988). Visual Perception. In R. Atkinson (Ed.), *Stevens' Handbook of Experimental Psychology*. New York: Wiley.

Hoffman, D. D. (2016). The Interface Theory of Perception. *Current Directions in Psychological Science*, *25*(3), 157–161. https://doi.org/10.1177/0963721416639702

Hoffman, D. D. (2018, March 23). The Interface Theory of Perception. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, pp. 1–24. https://doi.org/doi:10.1002/9781119170174.epcn216

Hoffman, D. D., & Singh, M. (2012). Computational evolutionary perception. *Perception*, *41*(9), 1073–1091. https://doi.org/10.1068/p7275

Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism: Basic Minds Without Content*. Boston, MA: MIT Press.

Isaac, A. M. C. (2014). Structural Realism for Secondary Qualities. *Erkenntnis*, *79*(3), 481–510. https://doi.org/10.1007/s10670-013-9519-3

Isaac, A. M. C. (2018). Prospects for timbre physicalism. *Philosophical Studies*, *175*(2), 503–529. https://doi.org/10.1007/s11098-017-0880-y

Isaac, A. M. C. (2019a). Epistemic Loops and Measurement Realism. *Philosophy of Science*, *86*(5), 930–941.

Isaac, A. M. C. (2019b). Realism without tears I: Müller's Doctrine of Specific Nerve Energies. *Studies in History and Philosophy of Science Part A*. https://doi.org/10.1016/j.shpsa.2019.01.002

Isaac, A. M. C. (2019c). Realism without tears II: The structuralist legacy of sensory physiology. *Studies in History and Philosophy of Science Part A*. https://doi.org/10.1016/J.SHPSA.2019.01.003

Isaac, A. M. C. (2019d). Realism without tears II: The structuralist legacy of sensory physiology. *Studies in History and Philosophy of Science Part A*. https://doi.org/10.1016/j.shpsa.2019.01.003

James, W. (1890). *The principles of psychology*.

Jones, L. A. (2003). Perceptual constancy and the perceived magnitude of muscle forces. *Experimental Brain Research*, *151*(2), 197–203. https://doi.org/10.1007/s00221-003-1434-4

Kelly, S. D. (2008). Content and Constancy: phenomenology, psychology, and the content of perception. *Philosophy and Phenomenological Research*, *76*(3), 682–690.

Kentridge, R. W., Cole, G. G., & Heywood, C. A. B. T.-P. in B. R. (2004). The primacy of chromatic edge processing in normal and cerebrally achromatopsic subjects. In *The roots of visual awareness: a festschrift in honour of Alan Cowey* (Vol. 144, pp. 161–169). https://doi.org/https://doi.org/10.1016/S0079-6123(03)14411-1

Kentridge, R. W., Heywood, C. A., & Cowey, A. (2004). Chromatic edges, surfaces and constancies in cerebral achromatopsia. *Neuropsychologia*, *42*(6), 821–830. https://doi.org/https://doi.org/10.1016/j.neuropsychologia.2003.11.002

Khan, S., & Chang, R. (2013). Anatomy of the vestibular system: a review. *NeuroRehabilitation*, *32*(3), 437–443.

Kiefte, M., & Kluender, K. R. (2008). Absorption of reliable spectral characteristics in auditory perception. *The Journal of the Acoustical Society of America*, *123*(1), 366–376. https://doi.org/10.1121/1.2804951

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt, Brace & World.

Kohler, W. (1947). *Gestalt Psychology*. New York: Liveright.

Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science Part A*, *29*(3).

Lee, D N, & Lishman, J. R. (1975). Visual proprioceptive control of stance. *Journal of Human Movement Studies*, *1*(2), 87–95.

Lee, David N. (2011). How movement is guided. Retrieved from http://www.pmarc.ed.ac.uk/ideas/pdf/HowMovtGuided100311.pdf

Lee, David N, & Reddish, P. E. (1981). Plummeting gannets: A paradigm of ecological optics. *Nature*, *293*(5830), 293–294. https://doi.org/10.1038/293293a0

Longino, H. E. (2013). *Studying Human Behavior: How Scientists Investigate Aggression and Sexuality*. University of Chicago Press.

Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. *Attention, Perception, and Psychophysics*, *80*(5), 1278–1289. https://doi.org/10.3758/s13414-017-1457-8

Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *Journal of the Optical Society of America A*, *3*(10), 1673–1683. https://doi.org/10.1364/JOSAA.3.001673

Maloney, L. T. (2003). Surface colour perception and environmental constraints. In R. Mausfeld & D. Heyer (Eds.), *Colour Perception – Mind and the physical world* (pp. 279–300). Oxford:

Oxford University Press (OUP).

Maloney, L. T., & Wandell, B. A. (1986). Color constancy: A method for recovering surface spectral reflectance. *Journal of the Optical Society of America, A, Optics, Image & Science*, Vol. 3, pp. 29–33. https://doi.org/10.1364/JOSAA.3.000029

Mark, J. T., Marion, B. B., & Hoffman, D. D. (2010). Natural selection and veridical perceptions. *Journal of Theoretical Biology*, *266*(4), 504–515. https://doi.org/10.1016/J.JTBI.2010.07.020

Marr, D. (1982). *Vision*. Cambridge, MA: MIT press.

Massimi, M. (2011). Structural Realism: A Neo-Kantian Perspective. In *Scientific structuralism* (Vol. 281, pp. 1–23). https://doi.org/10.1007/978-90-481-9597-8_1

Matherne, S. (2017). Merleau-Ponty on Style as the Key to Perceptual Presence and Constancy. *Journal of the History of Philosophy*, *55*(4), 693–727. https://doi.org/10.1353/hph.2017.0071

Matthen, M., & Rescorla, M. (2015). Bayesian Perceptual Psychology. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception*. https://doi.org/10.1093/oxfordhb/9780199600472.013.010

Merleau-Ponty, M. (1963). *The Structure of Behavior*. Beacon Press.

Merleau-Ponty, M. (1964). *The Primacy of Perception and Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics* (J. M. Edie & W. Cobb, Eds.). Northwestern University Press.

Merleau-Ponty, M. (2013). *Phenomenology of Perception* (T. Carman & D. Landes, Eds.). New York: Routledge.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press (OUP).

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*(3), 774–785. https://doi.org/10.1016/j.neuropsychologia.2007.10.005

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, *6*(1), 57–77. https://doi.org/10.1016/0166-4328(82)90081-X

Mitchell, S. (2003). *Biological Complexity and Integrative Pluralism*. https://doi.org/10.1017/CBO9780511802683

Mitchell, S. D. (2002). Integrative Pluralism. *Biology and Philosophy*, *17*(1), 55–70. https://doi.org/10.1023/A:1012990030867

Noë, A. (2012). *Varieties of Presence*. Cambridge, MA: Harvard University Press.

Peacocke, C. (1983). *Sense and Content: Experience, Thought, and Their Relations*.

https://doi.org/10.2307/2185393

Peters, M. A. K., Kentridge, R. W., Phillips, I., & Block, N. (2017). Does unconscious perception really exist? Continuing the ASSC20 debate. *Neuroscience of Consciousness*, *2017*(1). https://doi.org/10.1093/nc/nix015

Potochnik, A. (2017). *Idealization and the Aims of Science*. https://doi.org/10.7208/chicago/9780226507194.001.0001

Riley, M. A., Kuznetsov, N., & Bonnette, S. (2011). State-, parameter-, and graph-dynamics: Constraints and the distillation of postural control systems. *Science & Motricité*, (74), 5–18. https://doi.org/10.1051/sm/2011117

Rock, I. (1975). *Introduction to Perception*. New York: Macmillan Publishing co.

Rock, I. (1977). In defense of unconscious inference. In W. Epstein (Ed.), *Stability and Constancy in Visual Perception* (pp. 321–373). Wiley.

Russell, B. (1912). *The Problems of Philosophy*. London: Oxford University Press (OUP).

Schwitzgebel, E. (2008). The Unreliability of Naive Introspection. *The Philosophical Review*, *117*(2), 245–273. https://doi.org/10.1215/00318108-2007-037

Schwitzgebel, E. (2012). Introspection, What? In *Introspection and Consciousness*. https://doi.org/10.1093/acprof:oso/9780199744794.003.0001

Shockley, K., Carello, C., & Turvey, M. T. (2004). Metamers in the haptic perception of heaviness and moveableness. *Perception & Psychophysics*, *66*(5), 731–742. https://doi.org/10.3758/BF03194968

Smith, A. D. (2002). *The Problem of Perception*. Harvard University Press.

Smith, B. C. (2015). The Chemical Senses. In M. Matthen (Ed.), *The Oxford Handbook to Philosophy of Perception*.

Smith, B. C. (2017). Human Olfaction, Crossmodal Perception, and Consciousness. *Chemical Senses*, *42*(9), 793–795. https://doi.org/10.1093/chemse/bjx061

Worrall, J. (1989). Structural Realism: The Best of Both Worlds?*. *Dialectica*, *43*(1-2), 99–124. https://doi.org/10.1111/j.1746-8361.1989.tb00933.x

Wu, W. (2014). Against division: Consciousness, information and the visual streams. *Mind and Language*, *29*(4), 383–406. https://doi.org/10.1111/mila.12056

Zahorik, P., & Wightman, F. L. (2001). Loudness constancy with varying sound source distance. *Nature Neuroscience*, Vol. 4, pp. 78–83. https://doi.org/10.1038/82931

Zaidi, Q. (1998). Identification of illuminant and object colors: heuristic-based algorithms. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *15*(7), 1767–1776.

Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9656477

Zaidi, Qasim. (2002). Color constancy in a rough world. *Color Research & Application*, *26*(S1), S192–S200. https://doi.org/10.1002/1520-6378(2001)26:1+<::aid-col41>3.0.co;2-m