TESI DI DOTTORATO DI RICERCA IN CO-TUTELA

UNIVERSITÀ DEGLI STUDI DI MILANO
The Human Mind and its Explanations
Ciclo 36
Dipartimento di Filosofia Piero Martinetti

ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
Droit, Etudes Politiques et Philosophie
Institut Jean Nicod

As High as Eyes Can See:
A Moderate Liberalism for the Admissible Contents of Perception

Fin dove gli occhi possono vedere:
Un liberalismo moderato per i contenuti ammissibili della percezione

Aussi haut que les yeux peuvent voir :
Un libéralisme modéré pour les contenus admissibles de la perception

M-FIL/02 - Logica e Filosofia della Scienza

Quentin COUDRAY
Matr. R12853

Lavoro di tesi diretto da:
Corrado SINIGAGLIA (Università degli Studi di Milano)
Jérôme DOKIC (Ecole des Hautes Etudes en Sciences Sociales)

Coordinatore del dottorato:
Francesco GUALA (Università degli Studi di Milano)
Marc AYMES & Jérôme DOKIC (Ecole des Hautes Etudes en Sciences Sociales)

2023-2024

As High as Eyes Can See:

A Moderate Liberalism for the Admissible Contents of Perception

Fin dove gli occhi possono vedere:

Un liberalismo moderato per i contenuti ammissibili della percezione

Aussi haut que les yeux peuvent voir :

Un libéralisme modéré pour les contenus admissibles de la perception

Quentin Coudray

Ph.D. Candidate – *Università degli di Milano (The Human Mind and Its Explanations Doctoral Program) – Ecole des Hautes Etudes en Sciences Sociales (Institut Jean Nicod)*

UNIVERSITÀ DEGLI STUDI DI MILANO

L'ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES

Under the supervision of Jérôme Dokic (Ecole des Hautes Etudes en Sciences Sociales) & Corrado Sinigaglia (Università degli Studi di Milano)

Thesis conducted with generous support from Università Italo-Francese, Borsa Vinci 2021 (Capitolo II – Progetto C2-288)

UNIVERSITÉ FRANCO ITALIENNE | UNIVERSITÀ ITALO FRANCESE

CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my supervisors Jérôme Dokic and Corrado Sinigaglia for their support in the realization of this work. I would like to also dearly thank Francesco Guala for his support and his precious help in navigating the administrative labyrinths of the academic world. I also thank Vanessa Szenwald-Liwicki from EHESS for helping to set up the cotutelle agreement. Special thanks also to Università Italo-Francese for having granted me the Borsa Vinci 2021 (Capitolo II), and to the EHESS for having granted me a "bourse de cotutelle internationale", scholarships which allowed me to carry out my international doctoral work with some precious peace of mind. Finally, thanks to my colleagues, friends, family, and people from the Satellite for their relentless encouragements and love.

ABSTRACT

A philosophically crucial question within contemporary philosophy of perception is to determine what things we can perceive, as opposed to things we can only think about. In this thesis, I defend a "liberal" view of perception which accepts that we can perceive some kinds of high-level contents. I propose an original argument based on describing a relevant psychological mechanism that grants such representational capacity that I call *schematization*. Schematization describes a process by which perceptual systems (I focus on vision) representationally structure their sensory inputs, prioritizing certain feature dimensions, and implicitly activate (or prime) similar representations stored in perceptual memory. Schematization is a purely perceptual process that allows us to represent contents that are not reducible to low-level contents: *aspects*. Aspects represent some high-level *kind* properties of particulars. They represent particulars as having some physical body form that makes them belong to a superficial kind, such as the superficial kind of cat-form or chair-form. Crucially, I argue that aspects cannot represent natural or functional kind properties like cat-hood or chair-hood, since such properties depend on below-surface, non-visible characteristics of objects. I thus argue that careful empirical considerations about the representational capacities of perception vindicate a moderate Liberalism that only admits aspects representing superficial kind properties as the higher-level contents of perception. Aspects are as high as eyes can see.

ABSTRACT

Una questione cruciale nell'ambito della filosofia contemporanea della percezione consiste nel determinare ciò che possiamo percepire, in contrapposizione a ciò che possiamo solo pensare. In questa tesi, difendo una visione "liberale" della percezione che accetta che possiamo percepire alcuni tipi di contenuti di alto livello. L'argomentazione originale che propongo è basata sulla descrizione di un meccanismo psicologico rilevante che garantisce tale capacità rappresentazionale e che chiamo schematizzazione. Con questo termine si intende descrive un processo attraverso il quale i sistemi percettivi (mi concentro sulla visione) strutturano in modo rappresentazionale i loro input sensoriali, dando priorità ad alcune loro dimensioni, e attivando implicitamente (o innescando) rappresentazioni simili immagazzinate nella memoria percettiva. La schematizzazione è un processo puramente percettivo che ci permette di rappresentare contenuti, che chiamo aspetti, i quali non sono riducibili a contenuti di basso livello. Gli aspetti rappresentano determinate proprietà di alto livello degli oggetti, ovvero rappresentano gli oggetti come se avessero una forma fisica che li fa appartenere a un tipo superficiale, come ad esempio il tipo superficiale della forma-gatto o della forma-mela. Tuttavia, gli aspetti non possono rappresentare proprietà di tipo naturale o funzionale, come il fatto di essere un gatto o di essere una sedia, poiché tali proprietà dipendono da caratteristiche non visibili degli oggetti che si trovano sotto la loro superficie. Sostengo quindi che attente considerazioni empiriche sulle capacità rappresentazionali della percezione giustificano un liberalismo moderato che ammette come contenuti di livello superiore della percezione solo gli aspetti che rappresentano proprietà di tipo superficiale. Gli aspetti sono il limite fin dove gli occhi possono vedere.

RESUME

Une question philosophique cruciale dans la philosophie contemporaine de la perception est de déterminer quelles sont les choses que nous pouvons percevoir, par opposition aux choses auxquelles nous ne pouvons que penser. Dans cette thèse, je défends une vision "libérale" de la perception qui accepte que nous puissions percevoir certains types de contenus dits de haut niveau. Je propose un argument original basé sur la description d'un mécanisme psychologique pertinent qui confère une telle capacité de représentation que j'appelle la *schématisation*. La schématisation décrit un processus par lequel les systèmes perceptifs (je me concentre sur la vision) structurent de manière représentationnelle leurs entrées sensorielles, en donnant la priorité à certaines de leurs dimensions, et en activant implicitement (ou amorçant) des représentations similaires stockées dans la mémoire perceptive. La schématisation est un processus purement perceptif qui nous permet de représenter des contenus, que j'appelle les *aspects*, qui ne sont pas réductibles à des contenus dits de bas niveau. Les aspects représentent certaines propriétés de haut niveau des objets. Ils représentent les objets comme ayant une forme physique qui les fait appartenir à un *type superficiel*, tel que le type superficiel de la forme d'un chat ou d'une chaise. Il est essentiel de noter que les aspects ne peuvent pas représenter des propriétés de types naturel ou fonctionnel comme le fait d'être un chat ou d'être une chaise, car ces propriétés dépendent de caractéristiques non visibles des objets, situées sous leur surface. Je soutiens donc que des considérations empiriques minutieuses sur les capacités représentationnelles de la perception justifient un libéralisme modéré qui n'admet que ces aspects représentant des propriétés de type superficiel comme contenu de niveau supérieur de la perception. Les aspects sont aussi haut que les yeux peuvent voir.

Keywords: Perception, categorization, cognitive penetration, visual recognition, perception/cognition border

Mots clés : Perception, catégorisation, pénétration cognitive, reconnaissance visuelle, limite perception/cognition

Parole chiave : Percezione, categorizzazione, penetrazione cognitiva, riconoscimento visivo, confine percezione/cognizione

LIST OF FIGURES

I. INTRODUCTION

I.1. THE ADMISSIBLE CONTENTS OF PERCEPTION DEBATE: THE WHAT AND WHY

This thesis is about the *admissible contents of perception*. Perception puts us in contact with various properties of the world. It *represents* such properties. These perceptual representations have *contents*: represented *attributes* of *particulars*[1] in virtue of which such perceptual representations are *accurate* or *inaccurate*, depending on whether such particulars do or do not instantiate the properties represented by the attribute. I see a red cube. I thus perceptually represent a particular as having the following contents: the color attribute *red* and the shape attribute *cubical*. If the cube does instantiate the properties of *red color* and of *cubical shape*, then my perceptual representation is accurate.

The question of the *admissible contents of perception* concerns the thickness of the catalogue of contents that perception can represent. It concerns the types of properties that perception can attribute. Some contents are consensually accepted in the perceptual catalogue: it is not polemical that we can perceptually attribute colors, shapes, textures, sizes, distances, or motions (among others) to particulars in the environment. Many contents are much more polemical. Just to give an introductory illustration, one can wonder whether we can perceive a particular as being a *leopard*, beyond and above perceiving this particular's specific shape, texture, color, size etc.

This question is not trivial. We do trivially say such things as "I saw a leopard". But the way we talk about our perceptual states (and the way we talk of our mental states in general) is imbued with metaphorical usage. We also commonly hear things like "I see what you mean" or "I see that you passed your exam" (seeing someone's happy face). But those are clearly not literal statements about the contents of one's occurrent perceptual representation. Rather, they are simply statements of the fact that *from what one is perceiving*, one can *infer* what one is meaning or that one has passed her exam. These properties are not perceptually attributed but are inferred from perception and *thought about.* As Tyler

[1] I prefer to speak of the perception of "particulars" instead of the perception of "objects", because I will also accept that we can perceive liquids, shadows, mists, and spatiotemporal events. All these are perceived "particulars".

Burge[2] has aptly put it, "what ordinary language calls perception is a hybrid of perception proper and belief. Ordinary language blurs psychological kinds here" (Burge, 2010, p. 343, fn. 73).

This is to be contrasted with statements such as "I see a red cube": here, one is probably literally talking about the content of one's perceptual representation. One *perceptually* attributes the properties *red* and *cube* to a particular. Debates about the admissible contents of perception arise because there are many cases that live in some polemical grey zone. Seeing a leopard is one of them. One can legitimately ask: is it more like the case of "seeing that one passed her exam" or more like "seeing a red cube"? In other words, is it that we see some properties, and from them *infer* and *think* that there is a leopard; or is it that we *directly*, non-inferentially attribute the property "being a leopard", just inasmuch as we perceptually attribute color, shape, and texture? It is this type of question that is the heart of the admissible contents of perception discussion and that will constitute the overarching problem of this thesis.

In this discussion, one usually distinguishes between two types of contents. On the one hand, *low-level* contents are basic contents that are consensually recognized as being represented in perception. They include in vision[3] edges, orientations, sizes, distances, shapes, colors, textures, and motions (classically, but not exhaustively). On the other hand, *high-level* contents are more abstract contents that are *not* consensually recognized as being represented in perception. They might include (among others) *sortal* or *categorical* attributes ("being a leopard"), *evaluative* attributes ("being dangerous", "being beautiful"), *action* or *affordance* attributes ("being graspable"), or aesthetic attributes ("being harmonious"). It is contentious for each of these contents whether they are perceptually represented, or whether they are instead merely thought about upon perceiving some low-level properties.

Those who defend that perception can only have low-level contents are often designated as *perceptual Conservatives* defending perceptual Conservatism. Those who defend that perception can have *both* low-level and high-level contents are designated as *perceptual liberals* defending perceptual Liberalism[4]. We should note right away that there are many possible shades of perceptual Liberalism, since one might accept *some* high-level properties as being perceptually representable, e.g., accept that categorical attributes like "being a leopard" are representable, but maintain that others, like "being beautiful", aren't.

Finding a solution to the admissible contents of perception debate requires a uniting of philosophical and psychological reflections. Philosophy is needed because the debate trades heavily on conceptual distinctions between perception and cognition, and on getting a clear understanding of what high-level contents are. Psychology is needed

[2] Burge's philosophy of perception as developed in his two voluminous monographs (Burge, 2010, 2022) constitute the central philosophical framework which I redevelop in a concise form in this chapter.

[3] Like with most work in contemporary philosophy of perception, this work is almost exclusively dedicated to the *visual* modality. I think most of what I will say in this thesis can extend to other perceptual modalities

[4] For representative uses of the term, see e.g., (Bayne, 2009; Lyons, 2005; McClelland, 2016)

because the debate is ultimately one about psychological *capacity*, which is constrained by the human mind's computational setup and its underlying wetware. I think that neither philosophy nor psychology of perception can give on their own a satisfying answer to the admissible contents debate. Philosophical investigation alone seems to be bound to falter on introspective intuitions and question begging appeals to the "best" explanation for its defense of Liberalism or Conservativism. What it needs is to further base its arguments on empirically cogent psychological mechanisms. Psychological investigation alone runs the risk of an unreflective understanding of perception, since it will be understood instrumentally (so as to build experiments) as stimulus-dependence or activation in specific brain areas, while it should be considered a genuine *representational* kind in order to solve the targeted problem. Any proposed solution to the admissible contents debate on a purely psychological ground runs the risk of conceptual vagueness. Without philosophical work, we can't be sure that it is *perception* that psychological studies target, instead of some variegated amalgamation of different representational kinds. The debate should be about the admissible contents of *perceptual* states, not of gerrymandered perceptual-cognitive states.

The admissible contents of perception debate should take lessons from both these sources. This thesis aims to start to fulfil such ambition. It takes advantage first of immense theoretical leaps in contemporary philosophy of perception, especially in understanding the representational independence of perception from cognition (Block, 2023; Burge, 2022; Green, 2020a; Nes et al., 2023; Quilty-Dunn, 2017). It also takes advantage of the continuous progress in the scientific study of perception, which is one of the most advanced fields of research in the cognitive sciences. In particular, I think it is high time to reap the profits of the last decades of research on perceptual *recognition* or *categorization*, a research field that has been booming since the emergence of advanced brain imaging techniques and the development of deep neural network models that are directly relevant to the admissible contents of perception debate.

In this thesis, I will argue that such empirically informed philosophical theorizing should make one conclude in favor of a moderate kind of Liberalism. Conservatism will be castigated as false. But the Liberalism we should accept is not as radical as the one usually upheld in the literature (Bayne, 2009; Siegel, 2011). In particular, we should avoid thinking that perception can represent as "high" or "abstract" properties as the one that we can think about, such as *natural kind* properties (*cat-hood*), *functional kind* properties (*chair-hood*). I will defend a more moderate view, according to which perception represents its own garden-variety high-level properties: *superficial kind* properties. This will constitute the major positive thesis of this work.

A crucial underlying theme will be the distinction between the representational capacities of *cognition*, which are often conceptual and propositional, from the representational capacities of perception, which are non-conceptual and non-propositional. These differences in representational capacities explain why we can represent the natural kind *concept* "cat" in cognition, but not in perception. In perception though, I

still argue that we can perceive a cat-related high-level property, namely a *cat-form*. This is done through what I will designate as *aspect representations*. Aspects are purely perceptual representations. They correspond to a specific representational kind in the mind's economy that should be distinguished both from low-level perceptual contents (*contra* Conservatism) and high-level conceptual contents in thought (*contra* radical forms of Liberalism).

Why should it matter so much to determine what the contents of perception are? The general answer lies in the idea that perceptual contents form our most fundamental access to the external world. As Tyler Burge puts it, perception is "the first form of mind" by being the most primitive kind of representation of the world (Burge, 2010). By its representational nature, perception is expected to play a central role in the mind's economy, by being able to warrant beliefs (Burge, 2003), shepherd concept acquisition, or guide action (Nanay, 2013). *What* contents are perceptually representable will thus bear crucial weight on many corners of the philosophy of mind and epistemology.

A classical epistemological position is to claim that perceiving *x* (e.g., perceiving a light on the right) gives us *prima facie* epistemic warrant to believe that *x* (e.g., to believe that there is a light on the right). This view is often referred to as perceptual *dogmatism* in epistemology (Pryor, 2000). Dogmatism is both attractive as a folk epistemological theory and in professional philosophy. In our daily goings, we usually take our perceptual states as suitable justificatory grounds for knowledge claims. I see a keychain on my colleague's desk, and this is enough for me to claim to *know* that there is a keychain on the desk. If my colleague calls me and asks if her keychain is on the desk, I can confidently answer "yes", and if she feels epistemologically minded and further inquiries how I do *know* it, the answer "I can just *see* the keys on your desk right now" should be deemed enough.

In philosophy, dogmatism is also theoretically attractive. While philosophers are usually more cautious in claiming that perceiving *x* is sufficient to *know* that *x*, they have proposed (for different reasons) that perception provides *prima facie* justification (or warrant) to *believe* that *x* (Burge, 2003; Chudnoff, 2020; Goldman, 1976; Pryor, 2000). How one then gets from such *prima facie* justified beliefs to full-fledged knowledge is a matter of heated philosophical debate, but it won't concern us here. Dogmatism is philosophically attractive because it provides a partial response to skeptic arguments, in the sense that if dogmatism is true, then one need not accept the skeptic's claim that we have no satisfactorily warranted beliefs beyond reasonable doubt[5]. Perceptual beliefs can be taken to be satisfactorily warranted beliefs, barring defeaters (e.g., taking a hallucinogenic drug, suffering from a perceptual pathology etc.). Pryor usefully referred to those beliefs that are immediately justified by perception as "perceptually basic" beliefs (Pryor, 2000).

[5] As Pryor recognized (Pryor, 2000), perceptual dogmatism can only be a response to some kind of very strong form of skepticism that claims that *we* cannot have any prima facie justified beliefs beyond reasonable doubt. It cannot though counter the weaker skeptical argument that *truly* justified beliefs are impossible, because of the possibility evil demon scenarios.

An answer to the admissible contents of perception debate is of direct importance for the dogmatist epistemologist because it would tell her which beliefs are perceptually basic beliefs and which are not. Only under perceptual Liberalism will such beliefs as "this is a *Honda Accord*", "there is a *policeman* ahead", "the *cat is on* the *mat*" will count as basic beliefs, that is beliefs that are immediately epistemically warranted by what we are seeing. These beliefs would not rest on any further beliefs that would themselves require to be warranted. A belief epistemically relying on further beliefs for its justification can oppositely be said to be "mediately" justified.

Dogmatism by no means implies that *only* perceptually basic beliefs are appropriately warranted. Mediate beliefs which require further intermediate beliefs to be reached are also perfectly well warranted as long as intermediary beliefs are too. While under perceptual Liberalism we can accept:

1. I see a policeman ahead ["being a policeman" is a kind of high-level content].
2. Therefore, I am immediately warranted in believing that "there is a policeman ahead".

Under Conservatism we get:

1. I see a human-shaped body with a blue shirt ahead [perception only has low-level contents].
2. I justifiably believe that human-shaped body with blue shirts are usually policemen.
3. Therefore, I am mediately warranted in believing that "there is a policeman ahead".

In those simple examples, the final belief is equally well warranted under Liberalism and Conservatism. However, under Conservatism, the more "high-level" the belief to warrant is, the more intermediary steps will be needed to build an appropriate chain of warrant from the initial perceptual state. Raising the number of intermediary states raises the possibility of error, by raising the probability of introducing unjustified beliefs or invalid transitions. Of course, even the immediate epistemic warranting from perceptual states are *not* immune to error since perceptual states can themselves be inaccurate. But intermediary beliefs introduce additional possibilities of error by introducing other states, such as memories and beliefs that are also (if not more) prone to error. Thus, the category of perceptually basic beliefs plays an important epistemic role at least in the sense that they might be our most reliable (but fallible) beliefs, those that are the most immune to error.

Determining what contents perceptual states are able to represent also bears important consequences for debates relating to inter-individual differences in perception. The more liberal we are about perceptual contents, the more we can expect to find large inter-individual differences in perception. This might have troubling philosophical consequences, especially if one accepts a form of epistemological dogmatism. The main trouble is that one central characteristic of perceptual states is that they are hard to correct, or at least harder to correct than beliefs. One still perceives the two lines of the Müller-Lyer

illusion as being of different lengths even if one believes that they are of equal length. Perceptual disagreements are difficult to adjudicate if one cannot oppose a strong defeater to the competing perceptually basic belief. But precisely, we can expect that many such cases of perceptual disagreements will appear the more liberal we are about the perceptual catalogue of contents.

This is especially to be expected considering that the perception of high-level properties depends more intuitively on *learning* than the perception of low-level properties[6]. It is rare to disagree about the presence or absence of a low-level property. If I see that a pen is red, most other perceivers will too (barring, of course, they have no visual pathology such as colorblindness[7]). But once we consider higher-level, richer properties, disagreement is to be expected. Two foresters might *perceptually* disagree about whether a pine tree is a *Pinus pinaster* or a *Pinus pinea* (supposing Liberalism is true and that the foresters have learned to perceive subordinate kind properties of pine trees). These two foresters will have divergent basic beliefs. Solving such divergence is difficult if one does not get additional non-perceptual evidence.

Thus, Conservatism tends to limit cases of perceptual disagreement, because the perception of low-level properties is expected to be largely universal, whereas Liberalism tends to increase the potential for perceptual disagreements, by importing more idiosyncratic, learning-dependent contents in the realm of perceptual representations. The dogmatist who upholds Conservatism will thus make his epistemology rely on a perceptual ground floor that is mostly universally shared, whereas the dogmatist who upholds Liberalism will have to accept a perceptual ground floor that is replete with idiosyncratic divergences. Whether such idiosyncrasy of the perceptual ground floor under Liberalism threatens the epistemic warranting powers of perception is an important philosophical issue (Churchland, 1988; Fodor, 1984). I will come back to it in the very last concluding section of this thesis, when I will have defended my own version of Liberalism.

Finally, the Liberalism v. Conservatism debate might also bear some more surprising consequences for ethical questions. In fact, one important characteristic of perceptual states is that individuals are usually not taken to be *responsible* for having them. When one sees the two lines of the Müller-Lyer as being of different lengths, one is considered blameless for this misrepresentation. This is clearly different for cognitive states such as beliefs: if you believe something wrong (e.g., that Covid vaccines give cancer), then you can be taken to be blameful for this false belief. Humans are generally thought to have a moral duty to behave so as not to acquire false beliefs (in the constrained limits of the information they can have reasonable access to), but not to have a duty to not misperceive.

The answer to the Liberalism v. Conservatism debate will thus bear direct consequences for which mental states we deem the individual to be responsible for. The

[6] A genuine argument for this idea will be provided in VI. KNOWING HOW TO SCHEMATIZE.

[7] But in that case, colorblindness will uncontroversially be recognized as a defeater, so even the colorblind will usually not contest that the pen is red.

more we tend towards Liberalism, the more mental representations will fall outside of the individual's responsibility. This might have disquieting ethical consequences. Imagine we accept a Liberal view in which individuals perceive properties such as a person's *race* and a person's *threat-level*. Imagine that one individual systematically *perceives* Black individuals as more threatening than non-Black individuals (all other perceptual properties being equal). Should we conclude that this individual is not responsible and blameful of such racially biased perception, and of the basic beliefs that ensues? Things become even more disquieting when we consider behaviors that are driven by such perceptual representations. For instance, consider a police officer arresting a Black man and putting him in custody because he (wrongly) *saw* him as posing a serious threat. When the black man leaves custody free of charge, a trial is set against the policeman for a racial prejudice. In his defense, the police officer might explain his behavior by claiming: "I am absolutely sorry, but it is not my *fault*, I *saw* that man as posing a threat, and I can't control what I see!". In that situation, it might be delicate to blame the policeman for what he *saw* (of course, we can still blame him for what he *did*). Note that this kind of racial perceptual bias is not a philosopher's fancy, and cognitive scientists have claimed to demonstrate such perceptual effects in a growing literature about implicit racial biases (cf. Johnson & Wilson, 2019 for a recent review).

Now, this kind of "perceptual blamelessness" only arises if one thinks that all perceptual processes yielding representations of high-level properties are hardwired and outside of our voluntary control. It is such controllessness that grounds blamelessness. But it can be seriously doubted whether such absence of voluntary control is true of (high-level) perception. As will become clear in the thesis, I think Liberalism requires that one accepts that perception is importantly malleable and controllable in some ways. We can get an intuitive grasp of this idea with implicit racial biases: the most natural explanation of such biases is that they are *learned* in biased sociocultural environments (Munton, 2019), which are able to modify our mental states *down to the level of perception*. One might thus be blameful for his perceptual learning from biased environments.

Rejecting the hardwired view, one could thus claim that, just like for beliefs, one has a moral duty to not perceptually represent contents that are biased and inaccurate (again, in the limits of one's capacities). We thus end up with an opposite ethical conclusion than we began with. Because high-level contents are plausibly more sensitive to learning processes than low-level contents, the more one accepts liberal contents in perception, and the more one will have to accept a kind of *responsibility* for those perceptual contents. The unjustly arrested Black man could thus answer to the police officer that actually he *is* responsible for having *seen* him as threatening, since such perception of a high-level property was learned from an epistemologically inacceptable source (e.g. the police officer only watches bigoted news shows).

As captivating as these philosophical issues can be, they won't constitute the central target of this work. My aim here is only to point out that determining what kind of contents are represented in perception bears on many prominent philosophical issues, from

epistemology to ethics, and it is far from being merely a technical question within the philosophy of perception. I will come back to these wider philosophical consequences in the concluding sections of this thesis.

Having made this clear, it should also be quite conspicuous that the content of perception question is sufficiently intriguing on its own. After all, perception is our primary and most fundamental form of contact with the world. Thus, the question "what do we perceive?" is compelling enough on its own to deserve our philosophical effort.

The admissible contents of perception discussion cannot hope to progress if it is not made clear first what exactly is understood by "high-level", "perceptual", or "content". This is obviously true considering that all these technical terms are the subject of philosophical controversies. No sane discussion is possible if Liberals and Conservatives do not understand these terms in minimally common ways. Before starting the argumentative journey, I thus need to make explicit the best way to understand such terms. My first criterion will be that we should understand them so that the admissible contents of perception debate is not trivialized and remains philosophically palatable.

I.2. Seeing As Seeing-As

There are several understandings of "perception" that we must put aside to avoid trivializing the admissible contents of perception debate.

First, we should not take it as simply designating a causal relation between a subject and discriminable particulars in the environment, what we could designate as "discriminative perception", expressed in such propositions like:

S sees *o*

Where S is the perceiving subject and *o* the perceived particular. Claiming that S sees a leopard is uncontroversial if we are merely claiming that a particular in the world that we call a "leopard" is presently emitting some electromagnetic waves that are impinging on S's retinas, and to which S is discriminately sensitive. Discriminative seeing is equivalent to what Fred Dretske called *non-epistemic* seeing, where seeing is understood merely as visual differentiation of some particular (Dretske, 1969).

Discriminative vision is a "transparent" view of vision: one can uncontroversially replace the term "*o*" with any co-referential term *salva veritate*. If S discriminately sees a leopard, S can also be said to see a "*Panthera pardus*" or "John's favorite animal". All these latter propositions are all perfectly acceptable to the Conservative, even though they contain terms that represent high-level properties[8]. Thus, the admissible contents of

[8] I leave aside here issues regarding the ontological composition of the world. Conservatism and Liberalism are not theories about what ontology we should accept. A Conservative could accept that the world is composed of low-level properties *and* high-level properties like leopards but deny that the latter are

perception debate cannot take off if one understands perception as discriminative perception or "non-epistemic seeing".

Second, we should also not take perception in a mere *doxastic* or *epistemic* sense, as expressed in such propositions like:

S sees *that o* is a leopard

On this doxastic or epistemic understanding of perception (again, cf. Dretske, 1969 for a classical presentation), one "perceives that *p*" when one's justification for believing that *p* is wholly perceptual. As Alan Millar usefully puts it: *"*seeing that *p* is a matter of telling that *p*, prompted by current perception, and without conscious inference*"* (Millar, 2000, p. 77). When S sees that *o* is a leopard, it implies S has a belief to the extent that *o* is a leopard, a belief that is non-inferentially caused and justified by her current perception. Such non-inferential nature is intuitively guaranteed by the fact that S does not need to explicitly invoke any further beliefs to come to the belief that *o* is a leopard. Doxastic perception is obviously liberal: many beliefs about high-level contents are uniquely and non-inferentially justified by current perception.

A Conservative can perfectly accept these latter points. This shows that the doxastic understanding of perception is not the one at stake either. In fact, the Conservative might complain that this doxastic understanding overextends perception to post-perceptual *judgments* or *beliefs*, i.e., post-perceptual *cognitive* states that are only causally and epistemically related to perceptual states. The debate about perceptual contents should not be about perceptually grounded judgments or beliefs, since Conservatives admit that such representations can relate us to high-level contents.

This leaves us with a last understanding of perception, that of *seeing-as*:

S sees <u>*o as* a leopard</u>

I here introduce a crucial notational convention. I underline terms used to express mental contents. When S sees <u>*o* as a leopard</u>, S *perceptually* attributes the property *being a leopard* to the particular *o*. This example illustrates the fundamental representational nature of seeing-as. Seeing-as is always composed of two representational capacities (Burge, 2022, formalizes these capacities in detail): a *referential capacity* for referring to a particular *o*; an *attributional capacity* to attribute to such particular some specific properties (including kind properties like *being a leopard*). I follow Burge in accepting that in successful perception (excluding cases of hallucination), there cannot be perceptual reference without perceptual attribution,

---

representable in perception. Oppositely, a Liberal could accept that the world is composed only of low-level properties like shapes and colors but argue that perceptual systems represent high-level properties like leopards (in this case, these would be thought of as some kind of mind-dependent property). In any case, a Conservative can accept that we perceive *leopards* just in the sense that there is some particular in the world (denoted by the word "leopard", but denotable by other terms) that we are discriminately sensitive to. In this thesis,

or perceptual attribution without perceptual reference. One cannot perceptually refer to a particular without perceptually attributing to it certain properties: one cannot attribute properties without attributing them to a particular. In other words, and crucially, seeing-as is essentially dependent on discriminative perception.

When I write that S sees *o* as a leopard, this should be understood as S together applying a referential capacity to *o* and applying an attributional capacity to attribute the property *being a leopard*. In "S sees *o* as a leopard", the underlined part is the perceptual *content* of S's representation. It is what sets the accuracy condition for the perceptual state. If there is no particular referred to (there is no *o*), then the representation is inaccurate (it is a kind of perceptual hallucination). If the attributed property is not instantiated by the particular referred to (o is not a leopard), there the representation is also inaccurate (it is a kind of perceptual illusion).

Many times, I will often use the simpler phrasing "S sees a leopard", instead of "S sees o as a leopard". This is simply for reasons of stylistic fluency. Each time I use such phrasing with underlined terms, it should be read strictly under the seeing-as understanding, and *not* the seeing-that or discriminative seeing senses. Furthermore, if I write that one sees "a banana as yellow", this should be read strictly as one *discriminatively* seeing something (which I term a "banana", but I could have termed it differently with co-referential terms *salva veritate*), and attributing to it the property *yellow*. There is a crucial difference between "seeing a banana as yellow" (this is a low-level representation), and "seeing a yellow banana"(this is a high-level representation, because it includes the high-level kind content banana). What interests the admissible contents of perception debate is whether we can see something as a yellow banana, not whether we can see a banana as yellow (we obviously can).

When I write that "S sees a leopard" without underline, this because I am understanding seeing in a discriminative sense, i.e., I am *not* claiming that S attributes the property *being a leopard* to some particular (attribution which could be accurate or inaccurate). If I say that "S sees a leopard", it means S is able to visually discriminate some particular in the environment that I describe as being a leopard (but I could have written, *salva veritate*, that "S sees Mary's favorite animal", or that "S sees a *Panthera pardus*", since these are co-referential to the word "leopard").

Representational content is obviously not limited to perceptual representations. Thoughts, judgments, desires and beliefs also have representational content. Being faithful to my notational rule, I will thus write for doxastic perception:

S sees that '*o* is a leopard'.

I use apostrophes to emphasize the language-like nature of thoughts, i.e., most crucially, its predicative verb phrase syntax. Notice the difference with seeing-as, which

does *not* have a verb phrase syntax, but is expressed as a simple noun phrase (Burge, 2022)[9]. Perceptual judgments can anaphorically inherit referential applications of the perceptual representations to which they are bound. One can thus see *o* as yellow, black, furry, and anaphorically see *that* '*o* is a leopard, yellow, black, furry'.

Classically, one can think of thoughts (including perceptual judgments) as syntactic language-like compositions which compose over *concepts*, in addition to verbal connectors, logical connectors, or quantifiers (among others). I understand concepts as mental representations. I will from now on adopt the traditional notational convention for concepts and write them in small capitals. For perceptual *beliefs* or *judgments*, one should thus right:

S see that '*o* is a LEOPARD'.

Crucially, seeing *o* as *p* does not entail that S *believes* (or *judges*) that '*o* is *P*'. To get a better grip on the difference between seeing-that and seeing-as, one can observe that they do not attribute *error* (or *mis*representations) in the same way. If I see *that* 'the GAS TANK is FULL' by looking at the fuel gauge, but it turns out that the gas tank is empty (imagine that the fuel gauge is not functioning properly), then the error should be attributed to my judgment, *not* to my perception. There is nothing wrong here with my perception of the gauge (the needle *is* to the right of the gauge). What went wrong in this case is simply that my perceptually based judgment was unwarranted (to my ignorance)[10].

Contrast with a situation in which I see the gas tank *as* full, but it turns out I am undergoing a perceptual illusion, and the gas tank is actually empty. In that case, error should be attributed to the *perceptual* representation itself. Of course, one could *both* see the gas tank as full *and* see that 'it is FULL'. In that case, both representations would be

[9] Some authors have argued that perception is not syntactically structured, marking an essential difference with thought (Fodor, 2007). Other authors have argued that perception is syntactically structured, though not through a language-like verb phrase syntax, but rather through something like a proprietary *iconic*, *pictorial* syntax (Lande, In Press). I favor the latter idea, though nothing of what is to follows hinges on this commitment.

[10] Actually, this case is slightly more complex than it seems. In fact, the perceptual judgment is here *indirect* and derivative. One is not using the referential capacities of perception to make a perceptual judgment. The judgment does *not* refer anaphorically to the particular singled out by perception (here, the fuel gauge), but it refers to another *unseen* particular (the fuel tank). The judgment is thus not of the form '*o* is FULL' (where reference to *o* would be directly provided by perceptual referential capacities), but of the form 'the GAS TANK is FULL' (reference is guided purely by concepts). Such indirect judgment is still a good example of a perceptual judgment in some sense since its sole epistemic basis is one's current perceptual representation (though it might violate non-inferentiality). An example of a *direct* or *underivative* case of perceptual judgment being inaccurate while perception remains accurate is a case where one accurately sees the gauge as to the right, but inaccurately believes that the gauge 'is to the LEFT' (because maybe one does not master the concepts of LEFT and RIGHT). In this case nothing goes wrong in perception, but something goes wrong in perceptual judgment.

inaccurate. The important thing to note here is that seeing-as describes a representational state that is *perceptual*, while it is not the case for seeing-that.

Furthermore, seeing-that does not entail seeing-as. One could see that 'the GAS TANK is FULL' by looking at the fuel gauge, but this clearly does not entail that one sees the gas tank as full. In fact, seeing the gas tank as full requires that one sees some particular gas tank in the first place (in the discriminative sense described above). In other words, S seeing that '*x* is *P*' does not necessarily mean that S sees some *o* as *p*, since seeing that '*x* is *P*' does not even require that some *o* corresponding to *x* is discriminately seen, while seeing *o* as *p* *does* require that *o* is discriminately seen (i.e., does require discriminative perception). We can thus conclude that seeing-that and seeing-as are orthogonal representations (Overgaard, 2022).

This distinction further dovetails with recent empirical endeavors that aim at disentangling between *perception* proper (i.e., discriminative perception and seeing-as), and various kinds of perceptually-grounded beliefs and judgments. This endeavor is particularly important in the ambit of debates about how perception should be distinguished from cognition (Kanizsa, 1985). Cognitive scientists and philosophers have thus recently begun to look for diagnostic criteria of perception that could distinguish it from post-perceptual judgments (Block, 2023; Carranante, 2020; Firestone & Scholl, 2016a; Nes et al., 2023). Behind this thorny mission lies the premise that one should expunge seeing-that from our understanding of vision. Seeing is seeing-as, it is attributing some properties to particulars in a perceptual way.

This distinction between seeing-that and seeing-as is fundamental to make the admissible contents of perception debate consequential. In fact, if both are equated, then Liberalism would too easily come out as trivial. In fact, we can obviously believe many high-level contents from seeing something (that something 'is a PAINTING BY PICASSO', 'is a CAT', 'is a HAIRDRYER'…). If one accepts that perceptually-grounded beliefs are an integral part of perception, then obviously perception will contain many high-level contents. The seeing-that/seeing-as distinction thus should be clearly endorsed to make the contents of perception debate pertinent.

Still, this distinction is not consistently made in the philosophical literature. As Søren Overgaard notes, the tangling of seeing-that and seeing-as is quite common in many philosophical reflections on perception (Overgaard, 2022). It is illuminating to ponder on the reasons why such tangle of perception and perceptual judgments is so widespread in the philosophical literature, not least because it explains why the admissible contents of perception debate might be easily disregarded (and why many philosophers are Liberals by default). This tangling has its source in one influential view of perception according to which perception should be understood as *the subsuming of sensations (or sense-data) under some concepts or beliefs*. Here is for instance Thomas Reid:

> The external senses have a double province; to make us feel, and to make us perceive. They furnish us with a variety of sensations, some pleasant, others painful, and others

> indifferent; *at the same time they give us a conception*, and an invincible belief of the existence of external objects. (cited in: Buras, 2011, p. 6, *my emphasis*)

This "dual-component" or "doxastic" view of perception (Pelser, 2010; Quilty-Dunn, 2019) can be found in many other places throughout philosophy of mind. Here is for example Wilfrid Sellars interpreting Kant's theory of perceptual experience[11]:

> [Seeing] a cool juicy red apple (as a cool juicy red apple) is a matter of (a) *sensing-cum-imaging* a unified structure containing as aspects images of a volume of white, a sensed half-apple shaped shell of red, and an image or a volume of juiciness pervaded by a volume of white; (b) *conceptualizing* this unified sense-image structure as a cool juicy red apple. Notice that the proper and common sensible features enter in both by virtue of being actual features of the sense image structure and by virtue of being items conceptualized and believed in. As before, the applehood enters in only by virtue of being thought of (believed in). (Sellars, 1978, paragr. 24)

Under such views, distinguishing between seeing-as and seeing-that becomes arduous, since seeing is *defined* as the application of some belief or judgment to some sensations. On these views, seeing-as *just is* seeing-that.

Thus, doxastic views claim that:

seeing *o* as *p* = seeing that '*o* is P'

Notice that the doxastic view implies necessarily that perceptual tokens conceptual contents, since beliefs are made up of concepts (though the opposite is not necessary[12]). Even primitive demonstrative judgments ('*this* is a CAT') must employ concepts (that are predicated demonstratively). I strongly oppose assimilating perceptual representation to perceptual beliefs, and I further strongly oppose assimilating perceptual representation to conceptualization. Philosophers might often talk in the following way:

> We don't merely see the cat, we see her *as* a cat, a visual state that depends on conceptualizing her in a certain way. And in virtue of seeing her as a cat, we may come to form perceptual beliefs concerning her presence and qualities. Thus,

[11] Like Sellars himself insists, we should remain prudent and take this view to be first and foremost *Sellars'* view of perceptual experience that he thinks bears important resemblance with Kant's.

[12] A "non-doxastic" but still "conceptualist" view of perception could try to maintain a difference between seeing-that and seeing-as, by claiming that conceptualization of sensations (a kind of seeing-as) is purely non-propositional, while thoughts and judgments (seeing-that proper) are propositional (Weiskopf, 2015). I find such rejoinder unconvincing, since it remains to be shown how conceptualizing sensations could be effectuated without *predicating* some concepts to a sensory-subject, i.e., without forming a (primitive) proposition. There is a dilemma here: either conceptualizing is non-predicative (maybe it is purely demonstrative), but in that case one might legitimately wonder whether we are still talking about concepts at all, since these might reasonably be taken *by definition* to be mental representations that can enter into predications (Burge, 2009, 2022); or if conceptualizing is predicative, then seeing-as does really falls back to a kind of propositional seeing-that.

> conceptualized perception enables conceptualized thought. (Weiskopf, 2015, p. 223)[13]

I think this kind of talk is heavily misleading. Is Weiskopf talking about perceptual beliefs (seeing-*that* in my terminology)? In that case, the use of "seeing-as" terminology seems ill-advised, as it doesn't properly distinguish between perceptual representations proper (e.g., perceptually representing colors or shapes), and post-perceptual beliefs. Is he talking literally about perceptual representations? In that case, this amounts to taking it for granted that perception can represent *conceptually*, and that it can represent high-level contents. Such ideas shouldn't be taken as innocent presuppositions. They are heavily controversial. This thesis will argue that the former thesis (conceptual nature of perception) is not necessary for the latter to be (perception can represent high-level contents).

Why do many philosophers come to think that perception necessarily involves the subsumption of sensory experience under beliefs or judgments? I see two main reasons. The first is a phenomenological one. The idea that we do not only sense low-level sensory properties, but that we come to associate some abstract concepts with these sensations is taken to be phenomenally self-evident. We do not see only a smooth red round object, but we see an APPLE. Sellars again provides a good illustration of such an idea. In his article, he insists that his views only stem from "a store of accumulated knowledge" provided by phenomenologists. Thus, in a telling passage, he writes:

> The phenomenologist now asks us to take into account a phenomenon frequently noted, but as frequently misinterpreted. Consider the snow seen on a distant mountain. It looks cool. Do we *see* the whiteness of the snow, but only *believe in* its coolth. Perhaps this is sometimes so; but surely not always. Sometimes actual coolth is present in the experience, as was the *white* inside the apple and the *red* on the opposite side. Once again, we do not *see* the coolth of the snow, but we see the snow *as* cool; and we experience the actual coolth as we experience the actual whiteness of the snow. An actual coolness is bodily present in the experience as is an actual volume of white. (Sellars, 1978, paragr. 20)

The idea that we can *see* some things as having the high-level content COOLTH is deemed by Sellars so indisputable that he provides no specific argument for it. After all, we spontaneously believe something is cold when we see snow.

The second reason to accept a doxastic view of perception is epistemological. The idea is that if perception did not itself subsumed under some kind of conceptualization or belief, then it would put an impenetrable veil of images between perceivers and the world. It would just not be able to *represent* at all. Only sensations as coming under some beliefs are *about* the world. Sensations on their own are pure phenomenal events, they are not about anything. They have no intentionality. If perception consisted only of such

[13] What's to follow is not aimed at criticizing in particular Weiskopf's own views about (observational) concepts. I only take this passage as illustrative of a general tendency I oppose.

phenomenal sensations, then perception would fail to put us in contact with the world, it would fail to be *about* the world.

This is precisely what dual-component or doxastic theories are supposed to avoid (Pelser, 2010), by integrating within perception intentional states which are about the world and put us in contact with it. This epistemological impetus is for instance explicitly at work in the work of Thomas Reid, for whom his theory of perception was to be considered as a defense of a form of direct realism against any kind of solipsistic skepticism. For Reid (cf. the quote above), the most primitive kind of belief that is part and parcel of what it is to perceive is the belief in the existence of external objects, to which can be then added beliefs about the kinds of objects that are perceived ("conceptions"). Without such beliefs, we would be left isolated in our own world of sensations[14].

I think neither the phenomenological nor the epistemological arguments are strong enough so that doxastic views should be taken as the basic view of perception. In any case, in contemporary vision sciences and empirically-based philosophy of mind, such views stand as heterodox (Quilty-Dunn, 2019). These theories run into trouble too easily because they over-intellectualize perception. It is now widely recognized that creatures with very limited (if not completely absent) thought capacities can still be attributed with perceptual representations. They can still see things as *p* without *believing* that such thing 'is *P*'.

What's more, the epistemological impetus of doxastic and dual-component views is irrelevant in today's philosophical and scientific views of perception. In fact, the differentiation between non-intentional (non-representational) sensations on the one hand and intentional (representational) beliefs on the other has been largely abandoned in favor of a view where perception is viewed as a unified sensational-phenomenal *and* representational state in its own rights (though how these properties relate is still a matter of intense debates, cf. I.4). Perception is *in itself* intentional, and it doesn't need the application of beliefs or concepts to be about the external world. Equating "believing" and "representing" is heavily mistaken. Perception does have intentionality even when it is not subsumed under beliefs or concepts. So, we can safely put perceptually-based beliefs out of perception proper, i.e., distinguish between seeing-as and seeing-that, without sacrificing the intentional nature of perception. If perception is recognized as intentional, then it puts us in contact with the world, and can play a grounding epistemological role without being a kind of belief (Burge, 2003).

Finally, it just seems false to claim that perception systematically involves the application of some concept or belief even in creatures being conceptually competent. There are situations in which we can suspend our beliefs about what we see. This is particularly true in cases of known perceptual illusions: seeing the two lines of the Müller-Lyer as being of different lengths, I might still refrain from *judging* that 'the lines are of

[14] Tyler Burge argues that such view of perception that would be on its own unable to have representational powers (which he calls "Individual Representationalism") actually pervades much of the classic philosophy of perception (Burge, 2010).

DIFFERENT LENGTHS'. I might even judge something to be the case that *goes against* what I am seeing e.g., judge that 'the lines are of the SAME LENGTH'. Again, seeing *o* as *p* does not entail seeing that '*o* is *P*'.

The idea that it is supposedly phenomenologically indisputable that perception subsumes sensations under some beliefs strikes me as indeed disputable. True, we do spontaneously believe things about what we see, but spontaneity is just too weak a criterion to classify a mental state as perceptual. When asked "what's the capital of Italy?", I also spontaneously think and have the occurrent belief that 'it is ROME', but such thought is not a kind of perception (unless one is ready to understand "perception" in a very metaphorical way, but this is not what a philosophy of perception should do).

Also, it simply doesn't strike me as obvious that we see something as an APPLE CONTAINING A VOLUME OF WHITE JUICY FLESH. Imagine someone looking at a basket filled with real apples, next to a basket filled with visually indiscriminable plastic reproductions. Such subjects *knows* that the second basket is filled with plastic apples. In this case, it just doesn't strike me as obvious that one would perceive a difference between the two baskets. The much more intuitive thing to say here is that the two baskets as *perceptually* indistinguishable, even if one can have some beliefs to the effect that the basket on the left in filled with real apples, while the basket on the right is filled with plastic ones. I see no reason to think that these different beliefs should ground different perceptions, except on an unacceptable *petitio principii*. But I am not saying that such conceptualizations as of APPLES or JUICINESS are theoretically impossible in perception. I am only pointing to the fact that it is not a premise that we should take for granted. It is a position that must be argued for, and that is exactly what the admissible contents of perception debate is about.

We can thus conclude that the tangling of perceptual representations with perceptual *beliefs* is unwarranted. It can and must be premised that they are distinctive mental states, if only to make sense of the admissible contents of perception discussion. If we ask whether we can perceive high-level contents, what we are asking is whether one can one see *o* as *p*, where *p* is some high-level content. Liberals answer positively, Conservatives answer negatively. This is the central bone of contention of the admissible contents of perception debate, and it is a substantial philosophical question.

I.3. WHAT IT TAKES TO BE HIGH IN THE PERCEPTUAL ECONOMY

Next, we must clarify what exactly is meant by a *high-level content* or *high-level property*. First, I should make a reminder that the use of the adjectives "high-level" and "low-level" to talk of *properties* is only derivative. Strictly speaking, what is "low-level" and "high-level" are not worldly *properties*, but *mental representations*, and even more precisely the representational *contents* of those representations. Nonetheless, it is sometimes terminologically handy to also designate the properties that are represented as being "high" or "low" level. In this thesis, "high-level" properties, "high-level" contents, and "high-level"

representations will be used interchangeably. But one should keep in mind that ultimately, the debate between perceptual Liberalism and Conservatism is a debate about representational capacities, and not a debate regarding the ontological furniture of the world.

What exactly justifies designating some contents as *high-level* and some others as *low-level*? Most philosophers are just happy with an intuitive grasp of this distinction. Simply, low-level contents are those contents that are basic to perception and that are uncontroversially recognized as being perceivable (i.e., perceptually attributable). In vision, such contents usually include (among others): *edges*, *color* (including *hue, saturation, brightness*), *shape*, *texture*, *motion*, *size*. Conservatives defend that when a subject is seeing *o* as *p*, then how to fill in the *p*-place can only be done with one or a mix of such low-level contents.

Liberals argue that this is not enough. Fundamentally, they claim that one can see *o* as *p* where *p* is a content that is not reducible to one or any combination of low-level contents. The catalogue of such possible non-reducible contents in perception is extremely variegated and includes as diverse contents as:

- *Kind* or *sortal* contents, like *being a pine-tree*, *being a cat*, *being Joe Biden*, *being a natural scene* (Bayne, 2016; Calzavarini & Voltolini, 2022; Fish, 2013; Siegel, 2006, 2011)
- *Causal* contents like *is pushed by*, *is pulled by* (Butterfill, 2009; Michotte, 1963)
- *Affordance* contents, like *is sit-able*, *is grasp-able*, *is un-grasp-able* (Gibson, 1979; Nanay, 2011)
- *Social* and *mindreading* contents like *is angry*, *is chasing*, *is agreeable*, *expresses sadness*, *acts intentionally* (Block, 2014; Brogaard, 2016; Pacherie, 2005; Scholl & Gao, 2013)
- *Aesthetic* contents like *is beautiful*, *is an Impressionist painting* (Ransom, 2020c; Stokes, 2018b; Voltolini, 2023; Walton, 1970)
- *Moral* contents like *is cruel*, *is generous* (Audi, 2010; Cowan, 2015)
- *Ensemble* contents like the average size of a display's components (Bayne & McClelland, 2019)

This list is not exhaustive. Notice again that while it is enough to accept only one kind of high-level content (whatever it is) to count as a Liberal, a liberal position can be more or less strong depending on how many of these high-level contents it accepts. One could for instance be a narrow-reaching Liberal if one accepts that we can perceive affordances, but none of the other high-level contents. And one can be a wide-reaching Liberal if one accepts that we can perceive kind, causal, affordance, social, aesthetic, and moral contents altogether.

What exactly unites such a motley list? A tempting answer could be that high-level contents are constitutively contents that *supervene on*, are *grounded in*, or are *vertically articulated to* other lower-level contents. In other words, high-level contents would be contents that bear some causal dependence to other contents but that remain unreducible to them (Skrzypulec, 2018). Still put in other words, high-contents and low-level contents

would stand in a *determination* relation: low-level contents are *determinate* of *determinable* high-level contents. A pine tree is a determinable high-level content because it can be grounded in infinitely many *determinate* low-level contents (a pine tree can be represented as having many shapes, colors, sizes etc.).

While quite intuitive, I think such characterization is unsatisfactory. In fact, *most* if not *all* perceptual contents are dependent – or supervene on, or are grounded in, or are articulated to – *other* perceptual contents (or to non-representational sensations). Take organizational contents like contours or shapes. These are clearly dependent on perceiving other contents, like edges and orientations. A square is a determinable property, that can be grounded in infinitely many numbers of organizations of oriented edges. In Figure 1 below, the three objects are all seeable as square, though they are seen as having different orientations and edges.

Thus, being seen as a square is not reducible to being seen as having any specific arrangement of "lower-level" edge and orientations contents. Squareness should be considered a kind of high-level content if one took non-reducible dependence to be the relevant criterion. The issue is that the idea of perceptually representing something as square, or any other shape property for that matter, is not at issue in the admissible contents of perception literature. Even Conservatives accept that we can perceptually represent shape properties (which is often listed among the quintessential *low-level* properties). In other words, the criterion of non-reducible dependence overinflates the number of high-level contents. Worst, it might make *all* perceptual contents high-level in some ways. All organizational or "Gestalt" properties seem to precisely be defined by such non-reducible dependence on low-level properties (Wagemans, 2015). This is the case for such properties as symmetry, parallelism, groupings, or any particular Gestalts defined by spatial organizations (e.g., "face-gestalt", "cat-gestalt"). It doesn't seem warranted to focus the admissible contents of perception debate on these properties, since those are trivially recognized as being perceivable by vision philosophers and scientists alike.

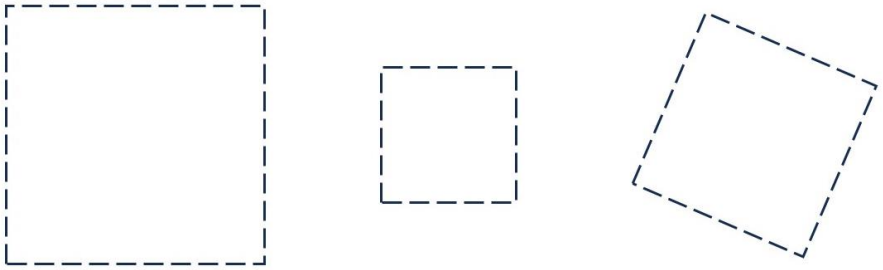

Figure 1. Three squares, each seen as square.

This issue actually extends to non-organizational properties that are also quintessential low-level properties. Take *size* as a first example. The perception of a size property (say, perceiving someone as 1,80m tall[15]) depends on taking into account other properties, most importantly *(egocentric) distance*. Thus, one can perceptually represent someone as being 1,80m tall when this person is viewed at an arm-length distance, or when viewed 30 meters away. When we are approaching someone, egocentric distance is reduced but our representation as of a 1,80m size remains constant. Thus, the size representation seems to be representationally dependent on (or be grounded in, be articulated to) representation of distance, while obviously not being reducible to it. We could say that size representations are *determinable* by distance representations. Seeing someone as 1,80m tall can be determined by infinitely many perceptions of distance. If we were to follow the intuitive definition of high-level contents as contents that are grounded in lower-level contents, we might have to count size as a high-level content. This is of course an unwelcome result: size is just a quintessential low-level perceptual content, so again it seems we shouldn't define high-level contents through representational dependence.

Just to give another example, *color*, yet again a paradigmatic low-level property, is also representationally dependent on other kinds of perceptual representations. One can perceive the same shade of red even when illumination conditions are dramatically changing. One actually *must* be sensitive to illumination conditions to perceive a particular shade of red, otherwise it would be impossible to perceptually differentiate between the color that is determined by illumination from colors that are determined by the surface reflectance of the object. Yet again, we could say that the same shade of red content is determinable by many illumination representations (imagine the same shade of red in the shadow or in bright sunlight)[16].

All the previous points are variations of the same crucial idea: *perceptually representing (attributing) a property is structurally dependent on perceptually representing (attributing) other properties.* The point has been made forcefully by Kevin Lande:

> Perceptual representations are perspectival because of the ways they are structured from their parts. What perspectival systems of representation have in common is that their representations of non-relational properties (such as size, shape, lightness, and color) and their representations of relational properties (such as distance, orientation, and illumination) are structurally interdependent. These structural interdependencies are rooted in the way perception works to disentangle the

[15] This is a simplification. I do not think that human visual systems have the computational capacities to attribute sizes in a precise metric way. Rather, they most probably attribute size in some less precise magnitudinal way.

[16] It could be that we never actually *perceive* illuminations (i.e., perceptually represent them), but that we are simply sensorily sensitive to it. Still, I think some kind of determination, grounding relation would persist. Color representation would be determinable by several *sensations* which make the perceptual experience of the color different.

> contributions that those properties make in producing sensory stimuli. (Lande, 2018, p. 214)

Notice that all properties Lande cites correspond to *low-level* contents, though they depend on other contents of perception to be representable.

These points show that it is a mistake to view perception as hierarchically divided between low-level contents that would be primitive and not dependent on any other representations (but only dependent on proximal stimulations) on the one hand; and high-level contents that would be grounded in such contents while being unreducible to them. Actually, perceptual contents are systematically computed in dependence to *other* contents. This feature is actually an *essential* feature of perception, because perception trades in the computation of perceptual *constancies* (Burge, 2010, 2022). Perceptual constancy is the ability to attribute a constant property *despite changes in other attributes of the particular or changes in sensory stimulations*. The proximal stimulations we receive are underdetermined: their sensory patterns can be the result of infinitely many distal causes. Information within proximal stimulation is tangled. As such, it is behaviorally unexploitable. For instance, it is underdetermined how much the size of a retinal projection is the result of distance and how much it is the result of the size of an object. Likewise, it is as such underdetermined how much the pattern of cell cones activation on the retina is due to illumination conditions, and how much is due to a particular's surface reflectance.

Amazingly, perceptual systems are able to solve such underdetermination through perceptual constancy mechanisms. Philosopher Peter Schulte has squarely exposed how this would work (Schulte, 2021). As we saw, the problem that perceptual constancy mechanisms must solve is that the proximal stimulation, let's call it K, carries information in a conflated form: it carries both information about a target variable, T, and information about a confounding variable, C. For instance, angle of retinal stimulation (K) carries both information about object size (T) and object distance (C). Size constancy mechanisms isolate the target variable T (object size), keeping it constant, and attribute changes in K (retinal angle) to changes in C (object distance). Consequently, we see the object getting closer, not getting bigger.

The isolation of the target variable from the confounding one is allowed by an auxiliary variable A, which tracks C independently of K. For instance, binocular disparity can play the role of an auxiliary variable in the tracking of object distances: the more an object is distant, the more there will be a difference between the right and left retinal images. When an object gets closer to a perceiver, binocular disparity decreases. The perceiver can thus establish, through changes in this auxiliary variable, that C is changing. Changes in K (retinal angle) can thus be attributed to C (the object gets closer), and T (object size) is kept constant.

The crucial thing to notice is that the computation of size constancy *requires representing (changing) distance*. Size is thus "perspectival" in Lande's sense.

Will Davies makes a similar point about color constancy:

> [An] organism's ability to achieve colour constancy supervenes on its capacities to discriminate visually among SSRs [surface spectral reflectances] of objects in the environment, and wavelength properties of the ambient illumination. (Davies, 2018, p. 663)

Again, color constancy is argued to be dependent on the capacity to represent another property, namely ambient illumination. Color is also "perspectival".

More generally, if one follows the steps of the influential account of Tyler Burge and accepts that *all* perceptual representations are constituted by constancy mechanisms, then *all* perceptual representations will be perspectival, *all* perceptual representations will depend on the representation of other properties[17]. High-level contents are thus not special in being hierarchically dependent on other contents. *All perceptual contents are*. The perceptual representation of a particular as a cat would be infinitely determinable by other contents (cats can surely be seen as having many distances-from-viewer, shapes, colors, textures, sizes etc.), but all other perceptual contents are *also* infinitely determinable.

This discussion is actually bound to a deeper point about perception linked to the idea that perceiving is perceiving-as: perceiving always comes in different *modes of perceptual presentation*. When one sees some particular as red, one sees it as red under a mode of representation. In the case of color, this mode of representation is dependent on *other representations*, such as representations of ambient illumination. One sees a particular as red in different ways when one sees it as in shadow or as in bright daylight. The point generalizes to all perspectival representations in perception. Saying that one sees *o* as red is thus incomplete, as it does not make explicit that there are many modes of representation under which such content can be represented. To be complete, we would have to say that one sees *o* as red^{p1}, where "p1" is a specific mode of representation (I borrow this formalization from Tyler Burge). Crucially, we should keep in mind that in many cases, perceptual modes of representation for a given content depend on other perceptual contents (though it might be that they also depend on non-representational sensations, cf. footnote 16). "p1" in this case depends on representing illumination conditions. Likewise, we should say that one can see something as 1,80m-tall^{p1} or as 1,80m-tall^{p2}, where the modes of representation p1 and p2 are dependent on distance representation for the referred to particular. This also works with shape-gestalts as in **Figure 1**: the three particulars are seen as square, though they are seen as square under different modes of perceptual presentation (square^{p1}, square^{p2}, square^{p3}). In this case, p1, p2, p3 are dependent on size and edge orientation.

[17] And *even if* one does not accept Burge's argument as to the constitutivity of constancy mechanisms for perception, and even if one accepts that there really are some perceptual representations that are primitive and are not dependent on other representations, this set of representations will still remain extremely limited (possibly to edges and their orientations). It would just be absurd and self-defeating for a Conservative to argue that only these primitive, "non-perspectival" contents are perceptually represented. It thus cannot be what a Conservative has in mind when talking of "low-level" contents.

This point is easily accepted for high-level contents like cat or pine tree: there are obviously infinitely many low-level contents to which these contents can potentially be articulated to, and these contents can have many modes of representation. But this point is not specific to high-level contents. Low-level contents are also articulated to other low-level contents and have many modes of representation.

But then what really does it take to be a high-level content in perception? I have actually come to believe that "high-level content" and "low-level content" do not form any natural representational kinds. At best, the distinction can only be understood as continuous. "High-level contents" are dependent on *more* other contents (they have more modes of representation, more "degrees of representational freedom") than low-level contents. Maybe "size" is low-level because it only depends on distance and edge representations. "Being a cat" is high-level because it depends on much more contents.

Is this vagueness in distinguishing between high- and low- level content an issue for the admissible contents debate? Clearly, *yes*. There are contents that are intuitively neither obviously high-level nor low-level. Organizational or "Gestalt" properties are the clearest examples. Some Liberals have argued that, since we can clearly perceive organizational properties (like face-organizations or face-gestalts both in human faces and in clouds), and since such properties are supposedly high-level (since they are multiply realizable, determinable and grounded in lower-level contents), then Liberalism is clearly true (cf. Calzavarini & Voltolini, 2022; Voltolini, 2023, for arguments in the vicinity).

The issue is that the criterion of non-reducible dependence does not clearly distinguish high-level from low-level properties. A Conservative can always answer such argument by claiming that organizational properties are not high-level (at least, the Conservative can legitimately claim that it is not *these* properties that she was denying in perception). Argumentative circles are clearly expected in such situations: Liberals will surely be tempted to accept such properties as high-level, while Conservatives will be tempted to think of them as low-level (or maybe "mid-level", see Landers, 2021).

What should we do? When there is no categorical definition, the best way to go is often to at least find some kind of *instrumental* understanding of a term. What understanding of "high-level content" would be useful? The first thing one can say is that it is useful if it makes the admissible contents of perception debate philosophically interesting and non-trivial. We should thus count as "high-level contents" only those perceptual contents that can be reasonably met with resistance by Conservatives, and that appeal for further theoretical and empirical exploration to determine whether they are really part and parcel of perception or not. I think this excludes organizational or Gestalt contents. It would be unreasonable for a Conservative to deny that such contents are perceptually represented.

Secondly, and probably most importantly, high-level contents are those contents that are not obviously present from introspective intuitions regarding perceptual phenomenology. The admissible contents of perception debate started out mostly as a debate over the contents of perceptual *experience* (Siegel, 2011). In perceptual

phenomenology, there are contents that clearly correspond to specific experiences: there is something it is like to see something as red, as a square, or as a straight line. But is there something specific it is like to see something as a cat, as a chair, or as Joe Biden, *keeping constant low-level contents and experiences to which they are articulated to*? Liberals answer positively, Conservatives answer negatively.

Introspective intuitions being infamously wobbly, high-level contents can simply be considered as those contents that are phenomenologically polemical to accept in perceptual experience[18]. Again, I think this criterion excludes organizational or gestalt contents. It seems phenomenologically pretty clear that such contents yield some specific perceptual experience. There is reasonably something it is like to see something as having a face-gestalt (think of the experience of suddenly seeing that a cloud looks like a face). Gestalt effects have a striking perceptual phenomenology (grouping effects; bistable figures; holistic perception etc.). Oppositely, it is not fully clear whether there is something perceptual it is like to see something as a cat (or as a CAT, if one accepts conceptual contents within perception) *beyond and above* seeing it as a specific cat-gestalt. A cat and a cat sculpture might share the same cat-gestalt, but is their a different experience corresponding to seeing a cat (the animal) and seeing a cat statue? This is an open question within the admissible contents of perception debate.

The list we established p. 16 does fulfil these minimal criteria for being high-level contents; they are contested by Conservatives, and they aren't phenomenologically obviously present in perceptual experience. In the remainder of this thesis, I will mainly focus on *kind* or *sortal* contents (as I think these are the high-level contents with the most chances of being perceivable). I will only still say some words about other potential high-level perceptual contents in the conclusion.

I.4. High-Level Perception, High-Level Perceptual Experience

If I see *o* as blue, there will usually be some phenomenal character attached to this perceptual state. I accept a roughly Fregean view of perception – i.e., I believe that perception always attributes properties *in specific ways*, or through specific *modes of representation.* I also accept that perceptual experience represents properties through different modes of (phenomenal) presentation[19]. One can visually experience *o* (say, a cube) as blue* under different phenomenal characters or different "perspectives": one can see *o* as blue$^{\star p1}$ when it is in

[18] Importantly, I warn that I do not think that the admissible contents of perception debate is uniquely about perceptual phenomenology (see the next section), though I do think that perceptual phenomenology acts as a crucial intuition pump in the debate.

[19] (Burge, 2022 Ch. 7; Chalmers, 2004, 2006; Sacchi, 2018; Thompson, 2007a, 2009, 2010) are all powerful defenses of this idea)

bright sunlight, and *o* as $\underline{\text{blue}}^{\star p2}$ when it is in darkness[20]. In both cases, one perceptually attributes the same property, but the conscious mode of representation of this perceptual attribution is different, thus yielding different perceptual experiences.

The phenomenal modes of representation of an attributed property are not unbounded and arbitrary though. One cannot attribute the property *blue* by way of the phenomenal character of the taste of chocolate. The point is that the phenomenal character constituting the mode of representation for a given property is fixed (certainly for an individual, and probably for a given species). For each represented property, there will be a range of specific phenomenal characters that are modes of (phenomenal) presentation of it. The limit on modes of representation for a given content is bound to how such content is structurally dependent on other contents for its representations (e.g., size contents are structurally dependent on various representations of distances; surface reflectance (color) contents are dependent on representations of ambient illumination).

If one argues that we can see something as instantiating a high-level property, it will thus be natural to argue that there will also be a range of specifically high-level phenomenal characters that are modes of (phenomenal) presentation of such property. In other words, there will be something specific it's like seeing something as instantiating a high-level property (Siegel, 2006, 2011). There is a proprietary *perceptual experience* of a high-level property, something we might call a "high-level phenomenal character".

The dispute over the contents of perception thus slides easily into a dispute over the characters of perceptual *experience*: are there some phenomenal character that are high-level, and that go beyond the phenomenal characters associated with low-level contents? Many presentations of the admissible contents of perception debate consider only this latter question (Price, 2009a; Siegel, 2011; Siewert, 1998). My thesis will diverge from these experience-first approaches. In fact, I will maintain a theoretical distinction between the admissible contents of perception debate on the one hand, and the admissible characters of perceptual experience on the other.

Obviously, such a distinction wouldn't make sense if one were to simply equate perception with perceptual experience. This is not the position I take in this thesis. I wish to follow common practice in cognitive science and accept that perception is a mental process that might or might not yield conscious experiences. I think there is good empirical evidence for the existence of *unconscious* perception, such as neuropsychological cases of blindsight (Taylor, 2020), or perception in very simple animals.

My own defense of perceptual Liberalism won't primarily rely on phenomenological considerations. It will instead focus on empirically grounded considerations about the kind of computations perception is able to perform on its proprietary representations. I think

[20] I use the superscript * to indicate that the perceptual content yields a particular perceptual experience with a specific phenomenal character. Such phenomenology is heavily dependent on mode of representation, so $\underline{\text{blue}}^{\star p1}$ and $\underline{\text{blue}}^{\star p2}$ are perceptual experiences that attribute the same property (*blue*), but that have different modes of representation (due to illumination) corresponding to different perceptual experiences.

these computations could be carried out even in a perceiving system that lacks any conscious perceptual experience, such as (potentially) some robots, animals, or phenomenal zombies. In other words, the Conservatism v. Liberalism debate about perceptual contents applies just as well for such systems: we can intelligently wonder whether a perceiving robot can perceive high-level properties or not (and just as well for animals for which we are not sure whether they have conscious experiences, and at the limit for phenomenal zombies).

The admissible contents of perception debate is thus relevant for any perceiving system which is deemed to possess perceptual representational capacities. Giving an answer to it for a specific system will greatly affect how we understand such a system's mental economy.

Furthermore, it can still be argued that perception safeguards its epistemological *warranting* powers even if it is not yielding any perceptual experience. This requires some further philosophical commitments, especially some kind of commitment to a form of epistemic reliabilism (Goldman, 1976). Perception can plausibly be viewed as prima facie warranting beliefs because perception is a reliable representational process, i.e., a process that normally, in some normal ecological conditions, produces accurate representations. Reliabilism does not require that a perceptual state be conscious to yield some epistemic warrant. A phenomenal zombie that *sees* a red cube and exclaims "a red cube!" would seem to be just as *prima facie* warranted in saying so than a conscious human perceptually experiencing a red cube exclaiming the same.

True, there has been dogmatist positions that are *phenomenal* dogmatist positions (e.g., (Chudnoff, 2011; Pryor, 2000). These positions argue that only conscious perceptual experiences can play an epistemically warranting role. The issue is that it is not clear what provides perceptual phenomenology such a privileged epistemic position. Why should seeing *o* as spherical (without phenomenal experience) be less epistemically warranting than *visually experiencing* *o* as spherical*? Seeing is representing a particular state of the world, even when such representation is unconscious. Why should such representation be less epistemically respectable than a conscious representation?

I think the larger philosophical background lying behind phenomenal dogmatism is some kind of *internalist* view of epistemic warrant: one is epistemically warranted just insofar as one has some *access* or *awareness* of the warranting basis. Phenomenally conscious states can be thought of as accessible states that the subject is aware of. The problem with this kind of rejoinder is that it is not made clear why a mental state being accessible is necessarily equivalent to one being consciously aware of this mental state. Even a phenomenal zombie could be said to have genuine access to his perceptual states (after all, he is only a *phenomenal* zombie, so it might be that the zombie still has something like *access* consciousness). So even on an internalist epistemology, it is not fully clear why it is only conscious phenomenal states that are epistemically warranting, because unconscious perceptual representations could also a priori play some evidential role.

"Reliabilist" dogmatists on the other hand have a straightforward story to tell as to why perceptual states (unconscious or conscious) have epistemic warranting force: perception is epistemically warranting because its representational processes are reliable. This reliability has nothing to do with whether such states are conscious or not (Berger et al., 2018).

Furthermore, the admissible contents of perception debate might also still bear important ethical consequences even if perception is unconscious. As we saw, attributions of responsibility for one's mental state intuitively depends on the level of rational control we can reasonably be expected to exert over such states. Whether one should be deemed responsible for her perceptual representations seems orthogonal to the question of whether such person has conscious perceptual experiences or not. Take the case of the racially biased policeman again, but now imagine that we are in a Philip K. Dick-like world where policemen are humanoid robots with human-like perceptual and cognitive capacities but no phenomenal experiences. In this world, it seems the robot-policeman could also defend itself against accusations of racial prejudice by claiming: "I am absolutely sorry, but it is not my *fault*, I *saw* that man as posing a threat, and I can't control what I see!" The rationale here is that the robot-policeman has not much control over what it sees, so that it should be blameless for what happened.

One could argue that genuine control over one's mental states is impossible without some kind of conscious experience, so that neither robots, animals, nor zombies should be considered responsible for any of their mental states (and probably any of their actions for that matter). But this is a heavy philosophical thesis, and one that cannot be taken for granted without some proper theorizing. It is reasonable to think that the burden of proof falls on such a position, since it seems to imply that control (and thus attribution of responsibility and rationality) depends on consciousness, an idea that is in itself surely quite controversial. As thing stands, it thus seems reasonable to think that the admissible contents of perception debate keeps with interesting ethical consequences even if we are dealing with perceptual representations that do not correspond to any phenomenal experience. Contrary to most approaches in the literature, my thesis will thus deal primarily with the admissible contents of *perception* debate, not the admissible characters of *perceptual experience* debate.

Whatever one thinks of these intricated epistemological debates, there is also a *methodological* point to be made in favor of approaching the admissible contents of perception debate not primarily from perceptual experience, but from perceptual representations, conscious or not. The point is simply that it is easier to rely on scientific constraints when talking of perceptual representations than when talking of perceptual experience. There is still as of today no consensual scientific way to understand perceptual experience. Introspection remains the principal investigative method in these territories. Introspection is infamously tendentious. The scientific study of perceptual *representations*, on the other hand, is a much more advanced field of study, maybe one of the most advanced fields of study in cognitive sciences. It thus seems a judicious choice to begin by exploring

first what are the admissible contents of *perception*, before considering what are the admissible characters of *perceptual experience*.

Now, having clarified this distinction, one shouldn't end up thinking that these two debates are isolated. Perception (understood as a mental-representational process, conscious or not) is deeply related to perceptual experience (understood as a phenomenal experience). What one *should not* end up thinking is that there is a direct and straightforward translation of the response to the admissible contents of perception debate to the response to the admissible characters of perceptual experience debate. How one should move from one to the other really depends on one's theory about the relationship between perceptual representations and perceptual experience. This issue is controversial.

On "Representationalist" theories of perceptual experience, the phenomenal property of one's experience is constituted by (and on stronger views, is *identical* with) its representational content understood as an attributed property (Dretske, 1995; Prinz, 2012; Siegel, 2011; Tye, 1995), or understood as a mode of representation of some attributed property[21] (Chalmers, 2006; Sacchi, 2018; Thompson, 2007a, 2010). In other words, Representationalism argues that mental representation is ontologically prior to phenomenology. This is supposed to be a decisive asset for Representationalism since one might hope that naturalizing mental representations will allow to naturalize phenomenology. Naturalizing mental representations is in itself philosophically challenging to say the least, but it is deemed by many philosophers to be in principle feasible (Dretske, 1995; Millikan, 1989; Neander, 2017; Shea, 2018), unlike the program of directly naturalizing phenomenology.

Under Representationalism, an answer to the admissible characters of experience debate might seem to depend on the answer to the admissible contents of perception debate. True, if a Representationalist is a Conservative about perception, then she must also be a Conservative about phenomenal characters, since under this view phenomenology is constituted by perceptual representations which are necessarily low-level. It seems there is a compelled coherence between the two answers.

Nonetheless, this is not true in the case of a Liberal Representationalist. In fact, it is possible that the specific subset of perceptual representations that she deems constitutive of perceptual experiences is a subset of states that only includes low-level representations. As E. J. Green well expresses:

> One might suggest, for instance, that the contents of representations produced by the visual system outstrip the contents of visual experience. If so, then even if a rich view were true of the visual system, it could be false of visual experience. (Green, 2017, p. 180, fn. 6)

It is theoretically possible for the Representationalist to be a Liberal about perceptual representations but to be a Conservative about the phenomenal character of perceptual

[21] I advocate for this latter "Fregean" understanding over the former "Russellian" understanding of perception in general. Perceptual content always come under fine-grained modes of representations.

experience. One could also be a Conservative about perceptual contents, but argue against traditional Representationalism that perceptual experience is also determined by factors beyond represented properties. For instance, Frédérique de Vignemont has argued that, while we might not represent *evaluative* properties in perception, we might still have specific *evaluative affective attitudes* towards some perceived properties, which itself fuses with perceptual experience (de Vignemont, 2023). This is a good example of a Liberal view of perceptual experience that doesn't have to assume Liberal perceptual contents (like evaluative contents).

Of course, a Representationalist might be a Liberal both about perceptual representations and phenomenal characters. This is for instance the archetypal Liberal position advocated by Susanna Siegel (Siegel, 2006, 2011; Siegel & Byrne, 2017). In that case, one will be said both to represent high-level properties (e.g., being a pine tree) and to have high-level phenomenal characters corresponding to such representation (e.g., a pine-tree* visual character), beyond and above the phenomenal characters of low-level representations (e.g., the shape* and color* phenomenal characters of the pine tree).

I am myself favoring a Representationalist approach to phenomenal character, through its "Fregean" understanding: I believe that perceptual phenomenal characters are grounded in some way in the *modes of representations* of perceptual contents (Sacchi, 2018). The particular phenomenal character of a specific shade of blue one represents depends on the mode of representation of this particular shade of blue. Under different illumination conditions of the same shade, this mode of representation will differ, and thus the perceptual phenomenology will differ, though the property perceptually attributed remains the same. I thus believe that if one manages to show that perception represents high-level contents and thus must have specific "high-level" modes of representation for them, then we can also conclude that there are some specifically "high-level" phenomenal characters. There will be something specific it's like to see something as a bat*, or to see something as a pine tree*. More correctly, there are *many* specific ways it's like to see something as a bat (bat$^{\star p1}$, bat$^{\star p2}$, bat$^{\star p3}$…) since there are many possible modes of representation for such high-level perceptual contents (just like for any perceptual contents). I think that phenomenologically experienced modes of representation bear non-trivial and necessary links to what content they serve to represent (*pace* Papineau, 2021). Thus, if one manages to determine the contents of perception, I think one will at least be able to circumscribe some specific modes of representation for them.

I thus think the defense of perceptual Liberalism that is to follow in this thesis can ultimately also serve to defend Liberalism about perceptual experience (high-level contents call for high-level modes of representation). But I do *not* think the argument of my thesis stands or fall on this latter conclusion. I think it remains philosophically possible and interesting to defend Liberalism *even if* such Liberalism is not bound to a Liberalism about

perceptual experience. My motto, like many these days in philosophy of perception is *representation first*, *phenomenology after*[22].

While the thesis focuses on the admissible contents of perception understood as a psychological-representational state instead of a conscious phenomenal state, this also does *not* mean that phenomenological considerations should be rejected. I do believe that perceptual phenomenology bears some intimate links with perceptual representations (cf. penultimate paragraph). Introspective intuitions stemming from perceptual experience can certainly help us in gaining insights into the nature of our perceptual representations. What I *do* reject though is an approach that would be *exclusively* stemming from introspective phenomenal considerations. I think we can and should do better to make progress on the admissible contents of perception debate by adopting a representation first, scientifically informed approach. This is what this thesis is doing.

I.5. Plan of the Thesis

The thesis is organized as follows.

In Chapter 2, I critically scrutinize previous arguments in favor of perceptual Liberalism. I argue that most of them are unsatisfactory on theoretical or empirical grounds. I argue that what is lacking is grounding vindication in favor of Liberalism in a consistent and plausible psychological mechanism.

In Chapter 3, I advance that research on *object recognition* is where to look for such mechanism. But I warn that such literature won't give a straight answer to the debate, since interpretative issues about whether such research program concerns *perception* proper and whether it really focuses on *high-level* contents or not are crippling clearheaded philosophical interpretation.

In Chapter 4, I show that however we think the object recognition literature should be interpreted, there is a psychological mechanism that is fit for grounding

[22] One of the main opponent of Representationalism, the "Phenomenal Intentionality" program (Horgan & Tienson, 2002; Pautz, 2009), claims that Representationalism has it all upside-down, because it is phenomenology that is prior to representation. Under Phenomenal Intentionality, we can attribute representational contents to a mental state only because such state has a certain phenomenal character, and not the other way around. For Phenomenal Intentionalists, asking what the contents of perception are without considering the character of perceptual phenomenology is thus nonsensical. On this view, one must first determine what are the admissible phenomenal characters of experience before determining the kind of representational contents such phenomenology can ground. My approach in this thesis is incompatible with the Phenomenal Intentionality program, unless what I say about perceptual-representational processes can be fully translated in non-representational terms. In that case, the order of explanation would be something like *non-representational physical processes* ground *phenomenal experiences* which themselves ground *representation*. I see nothing wrong with such position a priori, though I note that its explanation of the mental is in stark opposition to most scientific explanations in psychology, which surely gives a good reason to take such approach with a grain of salt.

perceptual high-level contents: *schematization*. Schematization is a process through which low-level perceptual representations prime similar perceptual memory traces under certain respects. I show that schematization is a purely perceptual mechanism and that it is empirically credible.

In CHAPTER 5, I consider the kind of contents that schematization produces: *aspects*. I argue that aspects are genuinely high-level contents (perceptual Conservatism is false), though they are not as "high" or abstract as often argued by Liberals. I argue that aspects can only represent *superficial* kind properties, i.e., membership into a group of particulars that is defined by *superficial* (or *body surface*) *similarity*. This is in stark opposition to most current forms of perceptual Liberalism which claim that we can perceive natural or functional kind properties.

In CHAPTER 6, I consider the relation between schematization and cognition. I argue that schematization is not *constitutively* dependent on top-down modulations from cognition. The purely perceptual nature of schematization is thus safeguarded. I show that schematization is learnable and modulable through different sources: evolutionarily hardwired selectivity, diachronic perceptual learning, and synchronic schematization control in cases of perceptual ambiguity. I consider whether this latter source of modulation should count as a kind of *cognitive penetration* of schematization (and thus of perception). I answer mostly negatively.

Conclusive CHAPTER 7 provides a general summary of the thesis, and opens up with general philosophical considerations. I advocate for some kind of "pluralism" about the mind's economy: the mind hosts many kinds of representations with different representational formats and contents. I conclude by reviewing the epistemological consequences of the moderate Liberalism defended in the thesis.

All in all, this thesis will be a success if I can show that Conservatism is false, but that we shouldn't adopt an overly radical version of Liberalism. Schematizing is not akin to thinking. Aspects' content in perception cannot be as abstract as concepts' content. Perception and cognition have their own representational realms, though these are in constant interactions.

II. How (not) to Defend Liberalism: Previous Attempts and Dead-Ends

How should we go about determining the admissible contents of perception? This chapter considers previous approaches, to prepare us for the introduction of my own original defense of Liberalism that will follow in the next chapters. Broadly, I think that the recent convergence of many different kinds of arguments in favor of Liberalism constitutes some good prima facie reason to believe that we can genuinely perceive some high-level properties. However, I think some of these arguments are simply non-starters, and I think that those that start from firmer grounds are still lacking a crucial argumentative ingredient. In fact, the liberal position is most often argued for in an *abductive* way, as being supposedly the best explanation for some intuitive or theoretical observation about perception. A recurring theme will be that such abductive arguments are as yet insufficiently convincing because Conservatives will be unconvinced that Liberalism constitutes the *best* explanation for such observations. As we'll see, Conservatives can always summon an alternative explanation that doesn't appeal to high-level contents in perception. To counter this argumentative conundrum, I will claim that we should add to such top-down *abductive* strategies a more bottom-up *inductive* strategy, one which defends Liberalism not just as the best explanation for some common observations, but that defends Liberalism as the best theory of perception's contents stemming from our current scientific understanding of perceptual processes.

In this chapter, I first consider four common defenses of Liberalism which I think are ultimately unsatisfactory (II.1 to II.4). These considerations though will allow us to clarify some important philosophical points about what a good defense of Liberalism should or should not look like. In the next part, I present a new kind of argument in favor of Liberalism that emerged recently: *common hallmark arguments* (II.5). I conclude by claiming that while such kind of argument constitutes an undeniable progress, they remain overly abductive in nature to be fully convincing (II.6). This chapter paves the way for a more positive defense of Liberalism that will be the topic of the following chapters.

II.1. Trusting what we Say or Think about what we See

For those who are not professionally involved in the admissible contents of perception debate as philosophers and psychologists can be, it is often difficult to grasp the tension between Conservatism and Liberalism (as I have learned across the years trying to explain my work to non-professional audiences). The main reason is that our folk psychology is naturally recklessly liberal. When we *talk* about our perceptual states, we naturally attribute to them the capacity to put us in contact with high-level properties. We say for instance "I *saw* Jack yesterday, and I could *see* he had missed his exam".

It should be clear that we shouldn't trust such ways of talking to get us ahead in the admissible contents debate. First, there are uses of perceptual verbs that are uncontroversially metaphorical, such as when we say, "I see a solution to the problem" or "I could see that my destiny was to stay here". Nothing properly *perceptual* is reported with them (even a blind person could utter such sentences).

Second, our use of perceptual verbs in our folk psychological reports does not make any distinction between the seeing-that and seeing-as understandings of perception (i.e., between its doxastic and representational interpretations). This is highly problematic since as I have argued only the latter "seeing-as" understanding is relevant to the admissible contents of perception debate. If someone says, "I see a leopard", is one reporting some kind of perceptual *belief* that she has while seeing an animal at the zoo or is one genuinely reporting on the contents of her perceptual state? Is one seeing the animal *as* a leopard? Or is one thinking that 'this is a LEOPARD'? Our folk psychology is unable to answer this question.

While the way we talk about perception is thus a dubious guide to determine its contents, some philosophers still argue that we have some kind of direct immediate introspective access to what perception represents. The common idea lying beneath these approaches is the claim that, instead of considering what we *say* about our perceptual states, we should consider what we *think* about our perceptual states. This idea is for instance advanced by Christopher Peacocke:

> We see things as tomatoes, and not as anything weaker. If the representational content of experience is given by what someone would judge, taking that experience at face value, then our ordinary experiences have a content concerning tomatoes, and not tomato-like objects. (Peacocke, 1983, p. 93)

What Peacocke means here by "taking experience at face value" is that we do not call upon further beliefs and background knowledge to make a judgment from our perception. I can judge that 'there is a CUP OF COFFEE on my DESK' at face value but judging that 'there is a CUP OF COFFEE from NICARAGUA on my DESK' requires to invoke some further beliefs, and thus cannot be a judgment made "at face value".

There are strong reasons to doubt that what we (could) judge at face value from our perceptual state is a good way to determine their contents (Price, 2009b). First, this position

is unable to make a needed distinction between the contents of perceptual judgments and the contents of perception proper. It is reasonable to expect that we can make perceptual judgments only based on our perception of low-level properties: we can judge that 'there is a TOMATO' simply because we see an object as having low-level diagnostic features, e.g., seeing something as tomato-shaped, red. This case is plausible, and the contents are clearly different. It might turn out that the object is not a tomato, but a plastic toy, so the perceptual *judgment* as to the presence of a tomato (the edible particular) is *false*, but the perceptual representation of a particular being tomato-shaped and red is *accurate.*

But on Peacocke's view, we cannot distinguish between these representations, since we can judge at face value that 'there is a TOMATO' from seeing something as tomato-shaped, red. Such view thus simply threatens to reduce perceptual representations to representations in perceptual judgments, and seeing-as to seeing-that.

Furthermore, Peacocke's method for determining perceptual contents only works insofar as we can introspectively break up judgements that are made at face values from those that are not. This premise seems dubious though since our judgments might rely on some hidden background knowledge form which we make some unconscious inference. My judgement seemingly "at face value" that 'this is a TOMATO' might rely on some background belief that 'red, tomato-shaped objects are TOMATOES', and not really be purely "at face value" after all. So Peacocke's introspective method seems unwarranted. I here join with a general critical point advanced by Jérôme Dokic that:

> [It] is not always transparent to naive perceivers what information is extraneous and what comes from experience itself. As a result, our intuitive self-ascriptions of experience (whether they follow Evans's procedure or Peacocke's more direct approach) are not infallible guides to what we perceive, strictly speaking. What we need is a firmer grip on the distinction between what is intrinsic to experience and what belongs to the cognitive background. (Dokic, 2010, p. 38)

David Lewis has proposed a somewhat similar view as Peacocke's, but instead of relying on our introspective access to "judgments at face value", he has argued that the contents of perception should be viewed as corresponding to those beliefs that are "ordinarily" produced by a perceptual experience. Thus, he claims:

> Visual experience depends on the scene before the eyes, and the subject's beliefs about that scene depend in turn partly on his visual experience. The content of the experience is, roughly, the content of the belief it tends to produce. […] We should take the range of prior states that actually exist among us, and ask what is common to the impact of a given visual experience on all these states. Only if a certain belief would be produced in almost every case may we take its content as part of the content of the visual experience. (Lewis, 1980, p. 240)

Lewis' method might seem superior to Peacocke's in that it doesn't presuppose we can know which judgments or beliefs are acquired at face value from perceptual experience. His method relies instead on the statistical regularity with which a belief is produced by a

perceptual experience. When we see a red tomato-shaped object, we tend to produce the belief that 'this is a TOMATO'. If most of us do, so the argument goes, it must be because the content 'is a TOMATO' is part of perception itself (whereas if this belief was dependent on other beliefs, which are supposedly more idiosyncratic, it would not be so universally shared).

But there are again strong reasons to doubt that Lewis' method is satisfactory in the ambit of the admissible content of perception debate. First, the view has trouble accounting for cases in which what we believe from our own eyes diverge between different people. Birdwatchers can recognize many different types of birds, while bird ignoramuses cannot. Should we thus conclude that nobody (not even birdwatchers) can perceptually represent subordinate types of birds, just because many in the population do not form any beliefs about them from seeing them? What about cases of recognizing persons, should we also exclude persons' identity as candidates for being perceived just because such representations will obviously not be universally shared? If Lewis' constraint really includes most of the population, then it seems it will just beg the question in favor of Conservatism, since only low-level properties are good candidates for yielding universally shared perceptual beliefs.

Another problem for Lewis' view is that it also seems to exclude from the perceptual representational catalogue properties that *do* seem obviously perceived. Imagine for instance a world where most of the population shares the background knowledge that the Müller-Lyer presentation is an illusion. In that world, anyone seeing a Müller-Lyer stimulus won't produce the belief that the two lines are of different length. On Lewis' account then, it seems we should conclude that the lines are not perceived as having different lengths. But this conclusion is highly problematic: it seems pretty obvious that we *do* perceive the lines as having different lengths.

Just like Peacocke's, Lewis' method thus seems unsatisfactory to guide us in determining the admissible contents of perception. More generally, *any* method that purports to determine the contents of perception by considering the kind of *cognitive* states it is spontaneously or ordinarily associated with (e.g., post-perceptual judgments or beliefs) will end up confusing the contents of perceptual beliefs (seeing-that) with the contents of perception (seeing-as). This is yet another marker of the problematic nature of any doxastic understanding of the contents of perception.

II.2. Phenomenal Contrasts

Instead of considering what we say or think from perception, another tempting strategy to determine the contents of perception is to consider its introspectable phenomenal characters. If one is a Representationalist about phenomenal characters – i.e., if phenomenal characters are equivalent to or at least supervene on representational contents –, then accepting that there exist some high-level phenomenal characters should make one conclude for the existence of some high-level visual contents.

How should one go about to prove the existence of such high-level phenomenal characters? The difficulty here is that we do not want to rely merely on introspective, individual intuitions about phenomenology, since we would probably just end up begging the question with Liberals arguing that it is obvious that there are high-level phenomenal characters, while Conservatives will resist such phenomenal intuitions (Dretske, 2015).

To avoid this conundrum, a popular strategy is to appeal to *phenomenal contrasts* arguments (Siegel, 2006, 2011). Phenomenal contrast arguments are thought experiments in which one is led to compare two perceptual experiences, which intuitively seem phenomenally different, but whose low-level phenomenal character is supposedly identical. In that case, so the argument goes, the phenomenal difference can only be explained by a difference in *high-level* phenomenal characters. Supposing that one is a Representationalist about phenomenology, then Liberalism would ensue.

To illustrate, take one perceptual experience $\underline{E_1}^\star$ such as seeing a car, and take another perceptual experience $\underline{E_2}^\star$, like seeing the same car and recognizing it is a Fiat 500 (imagine for instance that $\underline{E_1}^\star$ is my experience of a car when I don't know anything about Fiat cars, and $\underline{E_2}^\star$ is the experience I have after having lived 10 years in Italy and learning to recognize all types of Fiat cars). It seems a priori quite possible that the low-level phenomenal characters of $\underline{E_1}^\star$ and $\underline{E_2}^\star$ are identical[23] (the car hasn't changed color, size, shape etc.). Nonetheless, it is plausible that there is still something phenomenologically different between the two experiences. $\underline{E_2}^\star$ (recognizing a Fiat 500) visually doesn't feel like $\underline{E_1}^\star$ (simply seeing a car). From there, one can proceed to an inference to the best explanation to argue that the phenomenal difference between $\underline{E_1}^\star$ and $\underline{E_2}^\star$ is best explained by saying that they have different high-level phenomenal characters. And by Representationalism, this difference is explained by the fact that in $\underline{E_1}^\star$ one visually represents a <u>car</u>, while in $\underline{E_2}^\star$ one visually represents a <u>Fiat 500</u>. In other words:

$\underline{E_1}^\star$ = seeing <u>*o* as a car</u>$^{\star p1}$.
$\underline{E_2}^\star$ = seeing *o* <u>as a Fiat 500</u>$^{\star q1}$.

Since these different perceptual contents correspond necessarily to different classes of modes of representation (p*s* and d*s*), they will correspond to different perceptual experiences, thus explaining the phenomenal contrast[24].

Such phenomenal contrast arguments have been applied to defend Liberalism for many types of high-level contents, including *natural kind* contents such as seeing a <u>pine-tree</u> (Siegel, 2011), seeing <u>causation</u> (Siegel, 2011), or hearing the <u>meaning</u> of a sentence

[23] Of course, *modulo* visually experiencing the car from the same viewpoint, in the same orientation, under the same illumination conditions etc.

[24] This is my own Fregean understanding of phenomenal contrasts. Actually, Susanna Siegel's presentation of phenomenal contrasts is Russellian: she thinks the phenomenal contrast stems from a difference in the represented property *tour court*, not the mode of representation of such property.

(Siegel, 2006)[25]. Notice that the main purported strength of phenomenal contrast arguments is that the appeal to intuition is kept to a minimum: one needs only to intuitively agree that there is a phenomenal difference between $\underline{E_1}^\star$ and $\underline{E_2}^\star$, and then the argument springboards from there to an inference to the best explanation which is deemed to be free of a priori intuitions.

The issue though is that, even if a Conservative accepts that there really is a phenomenal contrast between two target experiences, she could always oppose that such contrast is *not* best explained by differences in high-level contents but might instead be explained in differences in low-level contents. So, the presupposition of the thought experiment that low-level phenomenal characters are kept constant might just be seen as begging the question in favor of Liberalism, and thus as not making any progress on the debate. This kind of critique is well put in a nutshell by Jesse Prinz's critique of Susanna Siegel's phenomenal contrast strategy:

> Even if she can show that attribution of k-properties [i.e. kind properties] can alter ordinary experiences, she cannot show by the method that the alteration results directly from k-properties, and not associated changes in attention, cross-modal associations, or imagery involving superficial visual features. (Prinz, 2013, p. 832)

So, for instance, the contrast between simply seeing a car and seeing a car and recognizing it as a Fiat 500 could be explained by the fact that in the latter case, one focuses her attention in a Fiat 500-specific way. In this case, one might attend to the car's headlights, with their Fiat 500-typical ovaloid shape, thus representing a new *low-level* content that one was not representing before like <u>ovaloid headlight shape</u>. Or maybe one spatially attends in a way that holistically embraces the specific curvy shape of the Fiat 500, thus making one perceive a new <u>Fiat 500 gestalt</u> organizational (low-level) content. Or still differently, when one sees a Fiat 500, one produces in inner speech a sentence like "this is a Fiat 500". In that case, the phenomenal contrast is explained purely in terms of non-visual phenomenology (it is explained either in terms of some auditory phenomenology or in terms of a proprietary cognitive phenomenology).

All these possibilities could explain the phenomenal contrast without appealing to a difference in visual high-level characters. Many Conservative authors have submitted criticisms of this kind, proposing alternative explanation of specific phenomenal contrasts that rely solely on differences in low-level perceptual characters or non-perceptual phenomenology (K. Connolly, 2014; Dokic & Martin, 2015; Dretske, 2015; Lycan, 2014; Nanay, 2011; Siegel & Byrne, 2017).

---

[25] Though most of these arguments do not usually take into *modes of representation* as determining phenomenology. Instead, they subscribe to a simpler Russellian Representationalist view under which what representationally differs between $\underline{E_1}^\star$ and $\underline{E_2}^\star$ is simply the attributed property (<u>car</u> vs. <u>Fiat 500</u>). I think this is a mistake, since on my Fregean Representationalism view attributed properties are not enough to determine phenomenology (one can attribute <u>blue</u> through many modes of representation).

Siegel is well aware of these potential alternative explanations, and she has offered extensive responses either preemptively (Siegel, 2011) or a posteriori (Siegel, 2013a). It is not my goal here to adjudicate whether phenomenal contrasts arguments can ultimately succeed or not. My goal is rather to point out that appeals to such arguments end up resting on polemical intuitions after all, since what is deemed to count as the *best* explanation for phenomenal contrasts seems to be notably theory-laden, thus raising doubts of question-begging abductive strategies. The inference to the best explanation in terms of high-level contents won't convince if explanations in terms of low-level contents or post-perceptual processes are just as much available and cogent to explain phenomenal contrasts. It would thus seem that Liberals must find cases of phenomenal contrasts where appeals to alternative explanation are not warranted.

Instead of phenomenal contrasts as given by thought experiments, a Liberal might want to strengthen her case by trying to look for phenomenal contrasts that can be empirically studied. For instance, Tim Bayne (2009) proposes that neuropsychological cases of *associative agnosia* confirm Liberalism. Associative agnosia is primarily characterized symptomatically as an inability to recognize objects. For example, in a famous case, one agnostic patient describes seeing "a long cord with a round thing at the end", but is unable to determine if it is a watch or a stethoscope (Rubens & Benson, 1971). Fundamentally, associative agnosia is thought not to impair the patient's capacity to perceive objects' basic low-level sensory qualities such as their shape and color. A reason to think so is that agnosic patients can copy accurately the objects they see, while they are unable to group these objects by categories.

If we accept the additional premise that it must *visually feel different* for visual agnosics to see objects compared to non-agnosics, then Tim Bayne claims we have a strong *empirical* case of phenomenal contrast that warrants a liberal conclusion. In fact, associative agnosia is commonly understood as an incapacity to *visually categorize* objects, or an incapacity to associate low-level contents to high-level categorical ("kind") representations. The contrast in visual phenomenology between associative agnosics and non-agnosics is thus straightforwardly explained by the liberal idea that non-agnosics have a visual experience of a high-level phenomenal property (corresponding to their categorization), while agnosics do not. While a non-agnosic experience a stethoscope*, the agnosic can only experience a long cord with a round thing at the end*. These content differences supposedly explain differences in perceptual phenomenology.

Does associative agnosia constitute a case for which the Conservative has no alternative explanation? Not really. Again, a possible alternative is to maintain that agnosics *do* have impairments in their perception of low-level properties, which would explain the phenomenal contrast without an appeal to high-level phenomenal characters. The assumption that agnosic patients have normal low-level perception relies mostly on the fact that they can produce accurate drawings of presented objects. But this assumption has been starkly criticized. In a review paper, Mette Kristine Hansen (M. K. Hansen, 2018), scrutinizes the empirical literature on associative agnosia and finds much counter-evidence

for this claim. First, while it is true that agnosic patients are reported to make accurate drawings of presented objects, they are also reported to be abnormally slow in their drawing, and to use abnormal drawing techniques such as drawing line-by-line instead of in continuous strokes. This might indicate that agnosics actually do have impaired low-level perception, and use "effortful strategies" in order to success in the copying task.

Second, Hansen notes that many neuropsychological descriptions of agnosic cases do report a deterioration in low-level perception. She sums up:

> Reports from patients suffering from agnosia indicate that their perception is not normal. Levine's patient reported blurriness in vision (D. N. Levine, 1978); Wapner's patient reported dimness (Wapner et al., 1978); and Rubens and Benson's patient reported that: 'faces seemed out of focus, almost as though a haze were in front of them' (Rubens & Benson, 1971). (M. K. Hansen, 2018, p. 309)

Some psychologists have additionally argued that associative agnosia results from a failure of integration of object parts into whole objects, instead of failure of object recognition. Going in this direction, Riddoch and Humphrey (1987) showed that agnosic patients focus abnormally on specific object parts for identification, disregarding the whole object. For instance, when and agnosic patient focused on an image of a baby carriage with metal spokes inside the wheels, she misidentified it as a bike, indicating that the patient might have only been able to focus on one part of the object (i.e., only seeing the object as having a wheel with spokes). Obviously, by representing only localized object parts, agnosic patients have trouble recognizing objects. Riddoch and Humphreys thus conclude that "agnosia may be determined by a specific deficit in integrating form information" (Riddoch & Humphreys, 1987a, p. 1431).

A further empirical observation supporting this idea is that agnosic patients are just as fast in copying impossible figures than in copying possible figures, whereas healthy patients are much slower copying impossible compared to possible figures (Ratcliff & Newcombe, 1982). This indicates again that agnosic patients might, contrary to healthy patients, only perceptually represent localized parts of objects (thus making them overlook the impossible-ness of the copied figure). Hansen concludes:

> These data indicate that the perceptual experiences of the agnosic patient differ from those of a healthy perceiver when it comes to which low-level properties these experiences represent, since they seem unable to see all parts of these simple objects simultaneously. (M. K. Hansen, 2018, p. 310)

These considerations show at least that Liberalism is not obviously the *best* explanation for cases of associative agnosia, even if one accepts that there really is a perceptual phenomenal contrast between agnosics' and non-agnosics' visual experiences.

Furthermore, a Conservative could also simply deny this premise. For instance, she could argue that the phenomenal difference is one in *cognitive* phenomenology: there is something *cognitive* it is like to recognize an object (objects might metacognitively feel *familiar*, cf. Dokic & Martin, 2015), and the agnosic patient might simply lack such

(meta)cognitive phenomenal character. Again, even considering empirical cases like cases of associative agnosia, it seems what the *best* explanation of them really is just falls back into the admissible contents of perception debate.

II.3. PSYCHOSEMANTICS

Instead of relying on intuitions from what we say, think or experience from perception, one might prefer to get one's hands philosophically dirty by directly setting up a semantics of perceptual representations. Under this psychosemantics strategy, one must "simply" choose a theory of mental contents and then apply it to perceptual representations to determine what their content is. If we agree about empirical facts about perceptual processes, and if we can reach agreement about a specific psychosemantics for perception, then it can be hoped that a common verdict will be reached about the contents of perception.

How can we determine the right psychosemantics for perceptual representations? It would be an understatement to say that building up a theory of mental semantics is a challenging and controversial enterprise. This might constitute a first overarching reason to be dubious about the chances of success of this approach: we don't want to end up defending a controversial position (either Liberalism or Conservatism) by an appeal to a psychosemantics theory that is even more controversial. This worry might however be alleviated by the fact that we do not need a semantics of mental contents *in general*[26], but merely a semantics for *perceptual* contents. By circumscribing the target of our theory of mental contents to perceptual states, we can hope to diminish disagreements.

Perception seems particularly well suited for a *causal* semantics, whereby "some syntactic item "X" means X because "X"s are caused by Xs" (Adams & Aizawa, 2021, sect. 3.). Simply, one sees something as a red cube because the tokening of such representation (the tokening of the representational syntactic vehicle corresponding to red cube) is caused by *red cubes*.

Such causal theory of perceptual content has tended to take advantage over the traditional "resemblance", "picture", or more contemporarily "structural isomorphism" (Cummins, 2000; Kulvicki, 2004; Shea, 2014) theories of perceptual content, whereby a mental item "X" means X because "X" *resembles* X. One perceives something as a red cube because one's perceptual state *resembles* or is *structurally isomorphic* to a *red cube* in a certain way.

While quite intuitive, resemblance theories have been pinpointed to face severe issues. First, the notion of "resemblance" is vague, and the structure of a perceptual state can resemble or be isomorphic with the structure of the world in a near infinite number of

[26] One might seriously doubt that it is even possible to provide a unified psychosemantics for all kinds of representational mental states, since different kinds of mental states (perceptual states, beliefs, desires, imaginary states, emotions etc.) might require their own unique psychosemantics.

ways, thus making the content of that state undetermined (Kulvicki, 2004). Consider a picture of Mick Jagger: this picture resembles at the same time Mick Jagger, a *Homo sapiens*, a living creature, an English man, etc. Resemblance theories thus seem to yield a strong indeterminacy for content individuation.

In addition, another crushing critique of resemblance theories is that they are unable to account for the "aboutness" of representational artifacts (like pictures) or mental representations. Explaining how intentional states are *about* or *refer* to the world is a required aspect of any semantics. But resemblance doesn't seem sufficient to provide this reference relation, mostly because resemblance can be obtained contingently. Take the phenomenon of pareidolia for instance: if a cloud resembles "the Sidney Opera House", it seems just wrong to assume that this cloud *represents* or *is about* the Sidney Opera House. Furthermore, resemblance doesn't seem necessary to obtain representational reference. If a kid draws anarchical doodles and claims that this is "her cat Larry", it seems to make sense to say that the drawing represents her cat Larry.

My aim is not to provide a full critique of resemblance theories of mental content, and I remain neutral as to whether some aspects of the theory might be integrated into a more general theory of perceptual content (Kulvicki, 2004; Ramsey, 2015; Shea, 2018). But I leave aside resemblance theories for now, to focus on the more conventional *causal* semantics of perception.

Causal semantics of perceptual content claim that one sees something as *p* if the tokening of *p* (more precisely, the tokening of the mental vehicle corresponding to *p*) is normally caused by *p* in the environment. We must instantly acknowledge a problem: just like resemblance theories, causal theories suffer from content indeterminacy threats. Consider the now classical case of a frog snapping its tongue at little black critters to feed itself (Schulte, 2012). Suppose a frog neural state *N* is activated by edibles in its environment, so that when *N* fires, the frog's tongue snaps toward the edible, resulting in its ingestion. But suppose that *N* is also activated by any small moving black thing (e.g., imagine an experimenter swings a little rock in front the frog). A question for the causal theorist is: what does *N* represent then? What property is it *normally causally* related to? Edibles or small moving black things (or something else)? As long as causal semantics haven't provided a clear solution to this "problem of indeterminacy" (Fodor, 1987), it will be impossible to leverage them to adjudicate between Liberalism and Conservatism about perception's contents. The issue, as we will now see, is that different versions of causal semantics end up with different verdicts on this question.

Teleological approaches, or "teleosemantics", determines the relevant causal relation of a perceptual state by determining the *function* such state is supposed to fulfil for the organism (Schulte & Neander, 2022). Thus, it is probable that the frog's neural state *N* represents edibles because the function of such state is ultimately to allow the frog to detect

edibles in its environment. The fact that *N* also responds to *small moving things* could thus be interpreted as a dysfunction yielding a misrepresentation as of edibles[27] (Dretske, 1986).

How might we go about determining the exact function of a mental state? Standard options include appeals to *evolutionary* functions: natural selection selects a mental state *N* only when its content participates in the greater adaptive fitness of the organism. To look for the function of *N*, one thus looks for the function that *N* representationally helps achieve that explains the enhancement of the organism's fitness. For instance, in the case of the frog, the biological function of neural state *N* is clearly to make the frog snap at *edibles*. It is snapping *edibles*, and not snapping *small moving black dots* that explains the participation of *N* to the fitness of the frog. Such evolutionary consideration of the biological function of *N* should thus make us conclude that *N* represents edibles, a kind of high-level content.

An alternative (and compatible) teleological approach is to postulate that a mental state's function can also be grounded *ontogenetically* through learning processes. Here, the frog's neural state *N* also represents edibles because, through reinforcement learning, the frog learns to eat *edibles*, not *small moving black things.* It is the edible representational content that is favored because learning to eat *edibles* enhances the organism's fitness, not learning to eat *small moving black things*. Postulating that the frog learns to perceive only small moving black things is not explanatory of its adapted behavior. Postulating that the frog learns to perceive edibles is explanatory of its adapted behavior.

Such teleological-functional approaches have been criticized on numerous sides (Adams & Aizawa, 2021; Burge, 2010; Schulte & Neander, 2022). It is not my aim here to assess whether teleosemantics is the good way to go for building a psychosemantics of perceptual contents. My goal is simply to suggest the following idea: teleological semantics of perceptual contents naturally lead to endorsing Liberalism. In fact, contents that will most often best explain successful behavior are high-level contents. Ever-changing and fine-grained low-level contents do not generally explain successful, adapted behavior. We behave because we supposedly see edibles, mates, or predators, and not simply cubes, red, or curvy things.

Of course, teleological theories do not deny that, at least at some point in the perceptual hierarchy, some perceptual states will represent low-level properties like small moving black things. Representing such properties is probably a necessary step for the organism to act adaptively, at least because representing those properties is necessary for the further representation of behaviorally relevant high-level contents like edibles. But a natural conclusion any teleosemantics seems to naturally arrive at is that, at some point, the frog must perceive edibles. If not, its tongue snapping behavior would be functionally and biologically unintelligible: why would a frog snap when seeing small moving black things?

[27] A teleosemanticist could further add that *false alarms* (seeing a small rock as an edible) are adaptively much less worrying that *misses* (failing to see an edible as an edible), so we shouldn't be surprised by the fact that the frog incorrectly represents any moving black dot as an edible.

This kind of functional-teleological argument for Liberalism can be generalized for many other high-level contents: it is probable that many animals should represent such high-level contents as predators, conspecifics, mates, food, poison, or shelter. To account for the adapted behavior of behaving organisms, teleosemantics thus naturally ends up advancing that perceptual representations should organize the world into behaviorally relevant high-level contents.

Does this constitute a satisfying argument for Liberalism? The main issue here is that teleosemantics is not the only approach to solve content indeterminacy, and competing alternatives can end up supporting Conservatism instead.

One main alternative to consider is what I would call *informational* semantics. Instead of solving content indeterminacy by considering which content best explains adapted behavior (by helping to perform specific biological functions), informational semantics solve content indeterminacy by trying to find the *fundamental cause* of mental state X, or the cause that is *informationally most powerful*. How exactly such notions of fundamental cause or informational power are elaborated will depend on the details of each informational semantics (Dretske, 1981; Fodor, 1987; Neander, 2017). But roughly, informational theories share the idea that to determine the fundamental cause of mental state X, say between potential causes c_1, c_2 or c_3, one should reason counterfactually and wonder for each cause, if c were absent, would mental state X still be tokened? E. J. Green nicely summarizes this idea by describing Karen Neander's causal-informational teleosemantics[28] (CT for short):

> CT incorporates a notion of property-sensitive causation, and Neander's most promising way of interpreting this notion enlists a screening-off principle. The latter principle tells us that to settle whether a representation R represents C or Q, where C and Q are coinstantiated properties of the external event e that caused R, we ask the following: (i) Would R still have occurred had e been C but not Q? And (ii) Would R not have occurred had e been Q but not C? If the answer to both questions is "yes," then C screens off Q, and so beats it out for the distinction of being R's content. (Green, 2017, p. 163)

We can apply this principle to the frog case. Does the frog perceive an edible thing or a moving black thing when neural state *N* is firing? The property of *being edible* is not the fundamental cause of *N* firing. In fact, this state is also tokened when *non-edible moving black things* are present. The property of being a *small*, *moving*, *black* thing *is* a fundamental cause of *N* firing. The firing state of the frog would *not* be triggered if the object were big,

[28] The terminology is rather confusing since Green classifies Neander's theory as subgenre of teleosemantics. This labelling issues should not concern us here though. What is important is to distinguish between semantic theories that rely mostly on downstream, output biological functions from those that rely mostly on upstream, input information to determine representational content. There are theories (like Neander's) that stand in-between.

immobile, or not black (as demonstrated by the fact that the frog does not snap at such objects), *even though a big immobile brown object might be edible for the frog*.

In the end, informational semantics of mental content arrive at an opposite conclusion compared to teleological theories regarding the frog case. While teleological theories favor a liberal conclusion (the frog sees edibles), informational theories favor a Conservative conclusion (the frog only sees small moving black things). In fact, we can expect that high-level contents will systematically be "screened off" by low-level contents in a causal explanation of the tokening of a perceptual state. Perceptual systems token the same states for stimuli that are exactly alike in terms of low-level properties, even if they differ in terms of high-level contents (a frog does not perceptually distinguish between an edible small moving black thing and a non-edible small moving black thing). The opposite is not true: if stimuli are exactly alike in terms of high-level property (they are both *edibles*), they might still be different in terms of low-level properties: compare an edible that is a *moving black thing* to an edible that is a *moving brown thing*. Only the former would trigger *N*. It seems then that the causally powerful or informationally fundamental property determining the tokening of *N* will systematically be the *low-level* property. It thus shouldn't be surprising that philosophers defending a version of causal-informational psychosemantics have mostly endorsed perceptual Conservatism (Dretske, 1981; Fodor, 1987; Neander, 2017).

In the end, the psychosemantics strategy has led us to some kind of *ignotum per ignotius*. How psychosemantics should solve content indeterminacy, and more precisely whether we should prefer the Liberalism-supporting functional-teleological views or the Conservatism-supporting causal-informational views is a philosophical matter just as contentious as the admissible content of perception debate we are trying to solve.

II.4. Cognitive Penetration

A tempting idea in the philosophy of mind can take the following form: if perception is *cognitively penetrated* by cognitive states, then Liberalism is true. Cognitive penetration of perception is said to occur when cognitive states like beliefs or desires synchronously modulate perceptual activity in a top-down way (Pylyshyn, 1999).

Why should cognitive penetration lead us to endorse Liberalism? The main reason is that, through cognitive penetration, the cognitive state would transfer its contents to perception. If the cognitive state has high-level contents, then the penetrated perceptual state will inherit this semantic richness. Many cognitive states such as beliefs and desires represent such high-level contents. In fact, cognitive states like thoughts are usually thoughts are usually considered to be syntactic compositions over *concepts*, and concepts can clearly represent high-level contents. We can have thoughts about CATS, JOE BIDEN, the CLIMB-ABILITY of a tree, the BEAUTY of a painting. If at least some of such high-level

conceptual representations of thought can penetrate and transfer to perception, Liberalism would be corroborated[29].

Should we believe in cognitive penetration? In a famous experiment by psychologists Jerome Bruner and Cecile Goodman (Bruner & Goodman, 1947), children are found to report that coins of greater monetary value have a bigger size compared to coins of lesser monetary value (the size of the coins in the experiment being kept constant). One natural interpretation of those results is in terms of cognitive penetration[30]: children have the belief that 'COINS are (more or less) VALUABLE objects', and this belief directly penetrates their perceptual state, making them perceive a VALUABLE COIN and not just a coin-shaped metallic object.

The (somewhat strange) premise of this argument is that seeing VALUABLE particulars makes them appear bigger. In that case, the interpretation of the case is straightforward: children see coins with greater value as bigger because they represent something that is VALUABLE in perception. VALUABLE COIN is obviously a high-level (conceptual) content, so Liberalism is vindicated by such case of cognitive penetration.

There are strong reasons to be doubtful about such defenses of Liberalism. Two central reasons can be invoked. First, the idea that cognitive penetration leads to a semantic transfer of the high-level conceptual contents of cognition in perception is far from being trivial. Nothing of the general definition of cognitive penetration requires this consequence. In fact, it could well be that cognitive states causally modulate perceptual states *without* transferring their proper high-level contents to perception. As Susanna Siegel (a Liberal) expresses it:

> The Rich Content View is not entailed by the thesis that visual experiences are cognitively penetrable. In principle, the influence of non-perceptual states on the contents of visual experience might be limited to non-kind [i.e. low-level] properties. (Siegel, 2011, p. 10, fn. 6)

Take Bruner and Goodman's coin experiment. It remains rather mysterious why representing a VALUABLE COIN in perception should lead to it being perceptually experienced as bigger. An answer here could be that valuable coins are perceived as bigger because children *attend* to those coins more compared to non-valuable coins. One thus adds a supplementary step in the causal chain from cognition to perception: children know that certain coins are more valuable than others, this in turns pushes them to attend more to valuable coins, such attentional command eventually modifying their perceptual experience of the valuable coin. The representation as of a VALUABLE coin-shaped metallic

[29] Paul Churchland has famously defended a radically liberal view of perception based on such strategy (Churchland, 1979, 1988).

[30] The results of this experiment, and more generally of many "New Look" psychology experiments that aimed at demonstrating cognitive influences on perception have been greatly contested (Firestone & Scholl, 2016b). For now, I consider these results as valid merely in order to present the cognitive penetration argument for Liberalism (critiques will soon follow).

object* thus produces an experience of a *bigger* coin that the mere (low-level) experience of a coin-shaped metallic object*.

The issue is that the introduction of such intermediary attentional mechanism threatens to break off the semantic transfer of high-level contents in perception. In fact, one could argue that attention can select, amplify, organize, or integrate the low-level contents of perception, but isn't able to generate new contents on its own. In other words, the view we get is one where a cognitive state with high-level contents triggers an attentional command which only modulates *low-level* contents. In that case, we should not conclude that the high-level representation VALUABLE is represented in perception. In this case, some beliefs merely causally modify the *low-level* contents in perception (i.e., the size of the coin).

Whether cognitive penetration implies genuine semantic transfers, and thus whether it implies Liberalism, will thus largely depend on how one defines cognitive penetration. But problematically, there is much disagreement in the literature on the topic. Some authors defend a wide view of cognitive penetration: any kind of synchronous top-down cognitive influence on perceptual activity will be considered a genuine case of cognitive penetration, be it through attention (Stokes, 2018a; Wu, 2017) or expectation (Lupyan, 2015). The issue for the defender of Liberalism is that such a weak kind of cognitive penetration does not carry with it the semantic transfers of high-level contents necessary to uphold Liberalism.

Other authors defend a narrower view of cognitive penetration where only those top-down cognitive effects which preserve "semantic coherence" with their consequential perceptual state are deemed to be genuine cases of cognitive penetration (Pylyshyn, 1999). On this narrow view, a causal modulation from cognition to perception isn't sufficient for cognitive penetration to occur. The modulated perceptual representation must in addition bear a certain semantic, logical or rational relation (I leave the relation intentionally vague) with its antecedent cognitive representation. As Steven Gross puts it:

> It is not enough that the contents of early vision states be sensitive to those of cognitive states in the weaker sense of depending counterfactually, statistically, or in a law-like manner upon them. They must also do so in virtue of early vision itself operating over the cognitive states *in a manner that mirrors a rational relation*. (Gross, 2017, p. 3)

What exactly is such a semantic relation? We can get an intuitive grasp of it through simple illustrations. For instance, in Bruner and Goodman's experiment, there doesn't seem to be any rational or logical relation between believing that 'COINS are VALUABLE' and perceiving them as bigger. The two mental representations are semantically unrelated, even if they are causally related (e.g., through attention). Another illustration of causal relation without semantic coherence is provided by Fiona MacPherson (Macpherson, 2012): imagine that Ben has been preparing for an important exam for several months, and is thus very stressed the evening before the fateful day. This important stress triggers a violent migraine, which

itself causes Ben to see flashes of light. Here again, there is no semantic coherence between the desire to pass the exam and perceiving flashes of light, even if there is clearly a causal relation between them[31].

Requiring that the modulated perceptual representation be semantically coherent with the penetrating cognitive state limits the scope of cases that will be considered as cases of cognitive penetration. This kind of restriction was already advocated by Zenon Pylyshyn to rebut the idea that top-down attentional modulations should count as cognitive penetration:

> Note that changes produced by shaping basic sensors, say by attenuating or enhancing the output of certain feature detectors (perhaps through focal attention), do not count as cognitive penetration because they do not alter the contents of perceptions in a way that is logically connected to the contents of beliefs, expectations, values, and so on, regardless of how the latter are arrived at. (Pylyshyn, 1999, p. 343)

Researchers like Zenon Pylyshyn or Steven Gross who defend the narrow view of cognitive penetration advocate that there are *no* uncontroversial cases of top-down cognitive modulation of perception that is "rational" or "semantically coherent". In fact, each time this kind of cognitive modulation goes through intermediary mechanisms like attention, semantic coherence will be thwarted. The high-level cognitive content is "lost in translation" so to speak and turns merely into a causal triggering force that modulates perception's low-level contents. So, while the issue for the wide view of cognitive penetration was that it couldn't support Liberalism, the issue with the narrow view is that it *would* support Liberalism, but it is so cognitively demanding that it remains highly controversial.

There is a potential candidate case that has particularly attracted the interest of philosophers and psychologists, which is that of "memory color effects" (Delk and Fillenbaum 1965; Hansen et al. 2006; for methodological critiques of this literature cf. Valenti and Firestone 2019). One talks of a memory color effect when the memory of one object's typical color (yellow bananas, red hearts, blue Smurfs etc.) affects the way people perceive those object's color. Usually, memory color effects result in subjects perceiving the object's color more towards their typical color. An orangish heart will thus be perceived as redder that an equivalent orangish square (Delk & Fillenbaum, 1965); a greyish banana will be perceived as slightly yellow (T. Hansen et al., 2006).

Fiona Macpherson (2012) has argued that the memory color effect is a case of cognitive penetration that meets the requirements of rational relation and semantic coherence between beliefs and perception. One sees a banana as more yellow than it really is because one believes that 'BANANAS are YELLOW'. Macpherson argues that memory color effects provide a superior argument for cognitive penetration because the effect is not going

[31] Note that a defender of the wide view of cognitive penetration could still argue that this isn't a proper case of penetration since the causal relation is diachronic instead of being online and synchronous.

through intermediary mechanisms like attention or expectation which thwart semantic penetration of cognition's conceptual contents in perception. She argues instead that memory color effects go through *imagination*: the belief that 'BANANAS are YELLOW' triggers the imagining as of a yellow banana, and this imagery states blends with and "paints over" perception, resulting in our perception of the grey banana as yellow.

For Macpherson, top-down modulation through imagination preserves the rational relation between beliefs and perception. The cognition-to-imagination transition is semantically preserving because imagination can be thought of as a process that translates cognitive content into a genuinely perceptual format (Kosslyn et al., 2006). Second, in the imagination-to-perception transition, the content of the imaginative state is confounded with the perceptual one, so that a unique perceptual state is generated. Macpherson argues that this sort of confounding of imagination and perception is pervasive in dreams and hallucinations. Furthermore, she claims that it has been experimentally demonstrated through the "Perky effect", where subjects confuse a faint projected image on a screen as being the result of their imagination (Perky, 1910). Thus, in the memory color effect, cognition modulates vision and does so in a semantically preserving way. One "translates" the belief that 'BANANAS are YELLOW' in cognition to a perceptual representation as of a yellow banana. Unlike for attention, imagination does not introduce an intermediary that breaks off semantic coherence between cognition and perception. Imagination preserves semantic coherence with high-level cognitive contents in its translation to perception (it might only transform a conceptual content to a nonconceptual equivalent). Macpherson concludes that memory color effects are genuine cases of cognitive penetration even considering its stricter, narrow interpretation.

Do we thus hold a good argument for Liberalism? Macpherson herself in her paper alludes to a strong relationship between narrowly understood cognitive penetration and Liberalism, claiming that "high-level content theorists are likely (though not compelled) to want it to be the case that cognitive penetration of perceptual experience can occur" (Macpherson, 2012, p. 33). Nonetheless, an argument for Liberalism based on memory color effect remains quite contestable.

First, it is only demonstrable for color contents, which are low-level contents and not at stake in the debate. One would need to demonstrate similar effects with *high-level* contents to uphold Liberalism through cognitive penetration.

Second, Steven Gross et al. (Gross et al., 2014) convincingly point out that Macpherson's explanation in terms of penetration-by-imagination is empirically problematic, for the time course of imagination is much slower than that of perception, which makes the effect implausible. In fact, if cognition was really modulating perception through imagination, we should observe a perceptual shift at some point. Perception is first uninfluenced by cognition because imagination takes time to be triggered. Then, when it is triggered, we should observe a color switch (e.g., from perceiving the banana as grey to the perceiving the banana as yellow). No such color switch is reported in the memory color effect literature, which threatens Macpherson's interpretation. Many authors have argued

that the effect can be instead explained by appealing to alternative cognitive mechanisms like decision biases (Zeimbekis, 2013), perceptual learning (Arstila, 2016), multimodal files (Deroy, 2013) or metacognitive feelings (Dokic & Martin, 2015).

Finally, some authors contest that memory color effects do exist altogether. Steven Gross et al. point out in their paper that the effect has issues of experimental replication. Chaz Firestone and Brian J. Scholl have argued in various papers that the literature suffers from various methodological issues (Firestone & Scholl, 2016a, 2015, 2016b; Valenti & Firestone, 2019). Some authors have eventually come to the conclusion that, to this day, there does not exist any empirical demonstration of cognitive penetration narrowly defined (Lammers et al., 2017).

To sum up, defending Liberalism through cognitive penetration, while prima facie tempting, is ultimately rather disappointing. In fact, this strategy leads to a troublesome dilemma: either one chooses a wide and encompassing definition of cognitive penetration, on which any cognitive top-down modulation of perception counts as cognitive penetration. While such effects are demonstrable empirically, they are in themselves insufficient to uphold Liberalism, because such effects aren't necessarily semantically preserving of the high-level contents of beliefs or desires that modulate perception. Or one can choose a narrower definition of cognitive penetration, on which top-down modulations must produce perceptual representations that are semantically coherent with cognition's representations. Such cases of cognitive penetration would support Liberalism. The issue though is that such cases of narrowly defined cognitive penetration are hard to demonstrate and are very controversial. The lesson to draw is not that such arguments are doomed to fail. Instead, it is simply that at best such defense of Liberalism might end up relying on premises that are just as controversial as the conclusion one wishes to defend[32].

II.5. Common Hallmarks Arguments

The ultimate approach to the admissible contents of perception debate that I wish to consider is a newcomer in the literature, a family of approaches which, following Madeleine Ransom, I will call "common hallmarks" or "common markers" arguments (Ransom, 2019).

The previous arguments I have considered (II.1 to II.4) focused on defending that some perceptual states can have high-level contents. That we were dealing with *perceptual* states was taken for granted. The common hallmark strategy defends Liberalism from an opposite side: it takes for granted some high-level contents and tries to determine whether those can be said to be perceptually represented or not. Thus, on this approach, the debate will mostly concentrate on determining criterions for classifying representations as

[32] The issue of cognitive penetration will come back later in this thesis, when after having upheld my specific defense of Liberalism, I will show that it does not uphold the idea of cognitive penetration understood in a narrow sense (VI.2.2 Schematization Control is not Cognitive Penetration).

perceptual, while on the other approach we focused on finding the semantics of states that were *presupposed* to be perceptual.

Why think this strategy will get us further? After all, the debate over the precise delimitation of perception seems just as contention as the one over how we should determine the semantics of perceptual representations (Carranante, 2020; Nes et al., 2023). Again, one might fear some vicious argumentative circles will ensue, with Conservatives choosing strict and narrow criteria for perception that exclude any high-level contents, while Liberals will tolerate looser criteria that encompass at least some of them.

But it's here that "common hallmarks arguments" can play their argumentative trump card: we are *not* compelled to have an exact theoretical delimitation of perception to build arguments as to whether a representation is perceptual or not. What we can do is to take a (high-level) representation and wonder if it bears some diagnostic *hallmarks* of representations that are consensually recognized to be perceptual by the scientific community. Thus, instead of a definitional strategy (does a high-level representation obey some definition for perception?), we adopt a less demanding *comparative* strategy (does a high-level representation resemble other perceptual representation?). In a nutshell, the strategy works as follows:

1. Choose a consensually *perceptual* representation *p*.
2. Verify the behavioral hallmarks that *p* exhibits when it is supposedly tokened.
3. Choose a high-level representation *h*.
4. Verify the behavioral hallmarks that *h* exhibits when it is supposedly tokened.
5. If the hallmarks of (4) are like those of (2), one can conclude that *h*, like *p*, is perceived.
6. Conclusion: Liberalism is true.

A central interest of this strategy compared to the previous ones is that its steps can be verified empirically through the methods of cognitive sciences. To fulfil step 1 and 2, we can look at representations of *low-level* properties which are consensually recognized to be perceivable. We can determine the behavioral hallmarks of perception by looking at what is typically observed when cognitive scientists study the perception of edge, contrast, color, motion etc. Of course, such hallmarks must be *diagnostic* of vision, so we will only retain those hallmarks that are uniquely exhibited by perception, and not exhibited by cognition (or other mental processes for that matter). For instance, being reportable is not a diagnostic hallmark of perception, since we can also report on our thought processes (through perceptual, seeing-that explicit judgements). As we'll see, there are many kinds of hallmarks of perception that can be devised, and different arguments will usually focus on a limited subset of them (if not a unique one). Overall though, we should expect that such hallmarks

will cluster together in some "homeostatic" way if we accept that perception is a natural mental kind (Boyd, 1991; Taylor, 2020)[33].

To fulfil steps 3 and 4, we can look at experimental tasks that, to be successfully completed, require the tokening of some high-level representation. Typically, such tasks will require subjects to classify stimuli into some high-level category. The cognitive science literature is replete with experiments studying our capacity to classify into high-level categories such as faces, animal species, car brands, or even artificial categories such as "Greebles" (Gauthier & Tarr, 1997). One can thus easily apply the common hallmark argument to such studies, determine whether subjects exhibit the hallmarks established in steps 1 and 2 when they complete them, and thus conclude, in an empirically grounded way, for or against Liberalism.

Common hallmarks arguments have philosophically boomed in recent years, overwhelmingly to the favor of Liberalism (Bayne, 2009, 2016; Bayne & McClelland, 2019; Block, 2014, 2023; Brogaard, 2016; Butterfill, 2009; Calzavarini & Voltolini, 2022; Cavedon-Taylor, 2021; Fish, 2013; Ransom, 2020a, 2020b). These approaches have considered various hallmarks of perception, but they converge toward the conclusion that high-level representations can exhibit perceptual hallmarks. This accumulation of evidence is a powerful asset for Liberals, for even if it turns out that some hallmark isn't found for some high-level content $\underline{h}$, or if it turns out that such hallmark is not diagnostic of perception after all, it wouldn't threaten the idea that $\underline{h}$ can be perceived if other diagnostic hallmarks are found. In fact, none of the hallmarks should be considered as *necessary* for considering a representation as perceptual. They are mere evidence for an abduction about a representation's perceptual nature.

What are the hallmarks of perception that we can observe experimentally and that are most often used by liberal philosophers in common marker arguments? As I said, the list is potentially very long. Let's take two recent representative examples to illustrate. Ned Block (2023) focuses on *adaptation aftereffects, binocular rivalry effects, pop-out effects, interpolation of illusory contours,* and *divisive normalization* (and he further cites but does not develop much: *sensitivity to Weber's law, search efficiency, innate functional role, interaction with exogenous attention,* and *perspectival sensitivity*). Madeleine Ransom (2019) on the other hand focuses on the following hallmarks: *adaptation aftereffects*, *automaticity, speed, parallel processing,* and *encapsulation*.

[33] I do *not* think that the admissible contents of perception debate necessarily rely on accepting that perception is a natural kind. Maybe perception forms some interesting cluster of properties, but not a natural kind in the same sense as "gold" is a natural kind. Maybe perception is merely a "scientific" kind that is useful to postulate for scientific discovery (Carranante, 2020). In that case, the admissible contents of perception would be about such scientific kind. It would not change the fact that knowing whether such scientifically interesting representational capacity characterized by such cluster of properties (as described by some hallmarks) can represent high-level contents or not is an interesting question. This non-naturally characterized representational capacity might still remain fundamental for epistemological, phenomenological, or computational issues.

In the next sections, I will quickly present Ransom's criteria which I deem to yield the most straightforward demonstration of how common hallmark arguments proceed from empirical data[34]. Converging results from these various common marker studies form the best evidence in favor of Liberalism to date, though we will soon see that some interpretative issues remain.

II.5.1. SPEED

When we perceive a low-level property, we seem to do so *instantly*. As soon as my eyes are open, I can see an object as cubical, blue, metallic. Of course, this intuition of instantaneity is misleading, since we know that perception takes time too (actually, neural processing is quite *slow* compared to that of, say, the electronic processing of a computer). What we can maintain is that perception is *faster* than cognition. Seeing an object as blue is faster than thinking 'this object is BLUE'.

Research on "ultra-rapid visual categorization" demonstrates that human participants can categorize objects into high-level kinds such as (e.g., animal, vehicle, bird) in the order of less than a hundred milliseconds (Grill-Spector & Kanwisher, 2005; Kirchner & Thorpe, 2006; VanRullen & Thorpe, 2001). Similarly, subjects can attribute high-level properties to scenes (e.g., *naturalness*, *openness*, *navigability*) at the same ultra-fast speeds (Greene & Oliva, 2009).

The idea that categorization can be very fast is corroborated by findings from electrophysiology. Intracranial neural recordings demonstrate specific neural response for representations of chair, fruit or animal in the temporal lobe as fast as 100 ms (Liu et al., 2009). Less invasive electroencephalographic studies also demonstrate specific event-related potentials (ERPs) for the representation of animals as fast as 150 ms (Thorpe et al., 1996). It has also been shown that the representation of faces generates specific ERPs as fast as 150 ms (Rousselet et al., 2008).

By the common hallmark argument, the fast speed of processing in tasks requiring discriminating some high-level contents is evidence that we *perceive* such contents.

[34] I will leave aside the encapsulation and parallel processing hallmarks. I deem the status of perception as being encapsulated and cognitively impenetrable to be too contentious and vague to serve as a good premise in the argument (cf. II.4). I deem "parallel processing" to be not specific enough to perception, since cognition (e.g., intuition) can also proceed in parallel and outside attentional focus (Chudnoff, 2020). This is not to be taken as claiming that these hallmarks are doomed to fail (maybe they could be refined). But my goal here is just to make a case as to the kind of abductive arguments that are possible from the most consensual hallmarks, and to scrutinize their limits.

II.5.2. Subtle Stimulus-dependence

An intuitive experimental hallmark of perception is that perception is stimulus-dependent while cognition isn't (Beck, 2018; Phillips, 2019). One cannot see a red cube with eyes closed, but one can think about a RED CUBE without a red cube being present.

Stimulus-dependence implies that there is a fine-grained causal mapping between physical properties of the stimulus and perceptual representations. This dependence is "fine-grained" in the sense that small and continuous changes in the stimulus parameters cause complex and unpredictable changes in the perceptual representations in ways that are difficult to verbalize for the perceiver. This sensitivity to small changes, as studied by psychophysics, is diagnostic of perception.

This provides us with an empirical protocol to test whether a representation *h* is stimulus-dependent in the relevant way: we test whether the completion of a task requiring representing *h* displays such kind of small-change sensitivity.

A direct illustration of this method can be found in the classical work from Belgian psychologist Albert Michotte on the perception of causal properties, which are instances of high-level properties (Michotte, 1963). Michotte studied the perception of launching relations between moving squares presented on screen displays. One of his central findings is that whether subjects report seeing one square as launching the other subtly depends on the interval delay between the latter and the former's movements. If this interval is too long, subjects do not report seeing launching but rather report seeing two objects move independently. Such dependence on the interval delay is extremely fine-grained: one only needs to raise the interval from 56 ms to 84 ms for subjects to perceive launching 100% of the time to merely 58% of the time, respectively. Representing launching is thus very finely dependent on small continuous changes in stimulus properties, a potential hallmark of its perceptual nature (Butterfill, 2009).

Stimulus-dependence also plays a central role in Brian Scholl and Tao Gao's (Scholl & Gao, 2013) argument that we can perceive mental properties in others, such as animacy (the property of intentional behavior). In one of their experimental designs, Scholl and collaborators show subjects displays with moving circles that are exactly alike in terms of low-level properties. One of these circles, "the sheep", is mouse-controlled by the participant. Another, "the wolf", is pursuing the sheep, and its pursuing behavior is parametrized by experimenters. Other circles are moving randomly. The subject's task is to detect which of the circles is the wolf, and to avoid being caught by it for a certain period.

Scholl and collaborators demonstrate that task success finely depends upon such parameters of the wolf's movements as its "chasing subtlety" (whether the wolf follows the sheep in a straight line or within the range of an angle) and its "chasing percentage" (the percentage of time that the wolf is chasing the sheep instead and moving randomly). Very small changes in chasing subtlety and chasing percentage parameters can dramatically change the success rate of participants in the task (i.e., the success at detecting the wolf).

This for the authors strongly buttresses the idea that we *perceptually* attribute animacy to a moving object. As they conclude:

> The point of reviewing these studies is to note that chasing detection (as a form of perceived animacy) is influenced in systematic ways by rather subtle display parameters, in the form of a psychophysical function (and in ways that do not seem readily explainable by appeal to higher-level judgment). (Scholl & Gao, 2013, p. 213).

By the common hallmark argument, subtle stimulus-dependence in tasks requiring discriminating some high-level contents like launching or animacy is evidence that we *perceive* such contents.

II.5.3. Automaticity

Another hallmark of perception is its *automatic* or *mandatory* nature (Fodor, 1983; Mandelbaum, 2015): once one opens her eyes, one cannot help but see. This is to be contrasted with cognitive processes like thought and reasoning, which are not thusly automatic and are deemed to depend on the subject's will.

How can one experimentally detect automaticity? A popular method is to build tasks whereby a representation *h* actually *impedes* completion of the task. For instance, *h* might be a distractor for detecting a target *t*. If one cannot help but represent *h* and cannot repeal this representation even if it hinders task completion, then we can conclude that *h* is *automatically* or *mandatorily* represented. For instance, if one must spot a blue square in a visual array, representing blue rectangles will stand as strong distractors, showing that the content rectangle is represented automatically (since it is still represented while it impedes task completion).

One can thus devise a task in which representing a high-level content would impede task completion. If this impeding representation is still tokened by the subject while completing the task, it is a good indication that it is *automatically* represented.

In a straightforward illustration of this idea, Alon Hafri and collaborators (Hafri et al., 2019) studied the automaticity of the representation of spatial relational properties, such as being in or being on. To do so, they presented rapid serial presentations of images that always contained two objects in some specific spatial relations (e.g., a knife *in* a cup). The subjects were tasked to detect the presence of specific objects by being presented with a test image containing two objects (e.g., a knife and a cup). Hafri et al. report that while subjects' overall accuracy is high (96%, Exp. 1), they tend to make errors significantly more often to distractor images *that display the same relational content as the target*. For example, when subjects must detect "a knife and a cup", they tend to make significantly more errors (false alarms) for "a pen *in* a basket" than for "a knife *on* a cup". This result as well as following control experiments support the idea that the perception of high-level relational properties, which in this case are irrelevant to the completion of the task, is automatic.

Another important experimental paradigm to demonstrate automatic representation of high-level contents concerns the representation of faces. A well-established characteristic of the representation of faces is its *holistic* nature: one does not see a face as an addition of contents (mouth, eyes, nose), but rather as an organized, unified whole, or face-gestalt[35]. It is a classic finding of the face processing literature that the holistic processing of faces occurs even when it undermines task performance. For instance, if a task requires subjects to determine whether the two upper halves of two face images are identical or not, the task is made more difficult if the irrelevant *bottom* halves of the face images are different (an effect known as the "Composite Face effect", cf. Murphy et al., 2017 for a review). The effect is strongly diminished or even disappears when the two halves are mis-aligned or when the pictures are presented upside-down. This shows that, when a picture is configured as a typical face picture, subjects automatically process it holistically, even if this undermines their capacity for task completion.

This kind of automatic holistic processing has actually been observed for many other high-level contents, such a face's gender (Zhao & Hayward, 2010), cars (Gauthier et al., 2003), fingerprints (Busey & Vanderkolk, 2005), dogs (Diamond & Carey, 1986), or artificial lab-created creatures like "Greebles" (Gauthier & Tarr, 1997).

These results suggest that a perceptual mechanism is at work in the representation of high-level contents since automaticity is considered a hallmark of perception.

II.5.4. Adaptation

When we perceive an object as having a certain property, our sensitivity to this property decreases, and the threshold for perceiving that property again is raised. This is referred to as *adaptation* aftereffects. For example, when one stares at a red surface for a certain amount of time, one's sensitivity to red decreases. Thus, if one then stares at a white surface (which mixes all visible wavelengths of the color spectrum), one will be desensitized to red, thus pushing the perception to the *opposite* of red, i.e., one will perceive the surface as being green. Adaptation produces *repulsive* aftereffects.

Do high-level contents exhibit such repulsive adaptation aftereffects? One major methodological issue for such hallmark (as noted by Burge, 2014) is that each time we will find a purported adaptation effect for a high-level content, one can suspect that what is actually adapting are actually the low-level contents to which the high-level content is dependent. One will thus have to be careful and control for such low-level confounds.

There are some empirical data that seem to fulfil this cautionary note. Ned Block (Block, 2014, 2023) makes use of research by psychologist Andrea Butler and collaborators

[35] A Conservative could legitimately complain here that holistic processing only shows that we perceive organizational or gestalt properties such as *face-gestalts*, so that it cannot weigh in favor of Liberalism after all. I think this point is warranted. It stems from the indeterminacy of what should count as genuine high-level contents (see I.3 What it Takes to Be High in the Perceptual Economy).

(Butler et al., 2008) demonstrating repulsive adaptation effects in the representation of facial expressions. In this study, Butler et al. made subjects stare at pictures of faces exhibiting facial emotional expressions (in this study: anger and fear). Then, they artificially created an ambiguous face stimulus by morphing faces with opposing expressions (fear + anger). They found that after having stared at one face expression (e.g., anger), subjects perceived the ambiguous face stimulus as exhibiting the opposite expression (e.g., fear), and *vice versa*. Crucially, authors took pain to control low-level confounds in different control experiments. Such controls were obtained by testing whether repulsive adaptation effects still occurred when facial expression were unrecognizable *but low-level properties were kept constant*. This was done by using hybrid faces (collages of facial expressions), scrambled pictures of faces, or scrambled line-drawings of faces. Authors found that adaptation aftereffects disappear (or are greatly diminished) in such cases, driving home the conclusion that these effects are really driven by the high-level facial expression contents.

Such repulsive aftereffects have been observed for other high-level facial contents, such as person's identity (Leopold et al., 2001), gender (Webster et al., 2004), race (*ibid.*), or age (Schweinberger et al., 2010). Even more strikingly, one study (Javadi & Wee, 2012) demonstrated that adaptation aftereffects can be obtained for gender when one is adapted to a gender-typical *object* (e.g., lipstick for women, electric shaver for men), and then is shown an ambiguously gendered face. For instance, when one is adapted to a picture of a lipstick, one will tend to perceive a picture of an ambiguous face as being masculine. This is striking because in this case there is no direct adaptation to some low-level content that could explain the repulsive adaptation effect. How could the low-level contents of a lipstick picture desensitize the low-level contents of a face picture, considering that they are so different[36]?

Another study has even demonstrated a *cross-modal* adaptation aftereffect for facial expressions (Matsumiya, 2013): haptic exploration of a face mask (hidden frow view) which expresses happiness makes one *see* a neutral face as more sad (and *vice versa*). Here again, we cannot appeal to adaptation to low-level features driving the effect, since these features are not even in the same modality (thought see fn. 26 again).

Most of these studies on high-level adaptation concern face-related contents, and there are very few studies that have reported on such effects for other high-level contents. An interesting exception is Michelle R. Greene and Aude Oliva's demonstration of high-level adaptation effects for scene-specific contents such as naturalness, navigability or even their temperature[37] (Greene & Oliva, 2010). This result is even more interesting considering that these scientists also demonstrated *ultra-fast* representation of such scene contents like naturalness (Greene & Oliva, 2009); a finding of a conjunction of hallmarks (speed and

[36] But it remains to be determined whether the gender-typical pictures do not trigger at least some mental *imagery* of a gendered face whose low-level contents might explain the adaptation effect.

[37] Was Sellars right after all in thinking that one can see the coolth of snow?

adaptation) for a unique content (naturalness) that is exemplary of how common hallmark arguments can be virtuously combined.

Demonstration of repulsive adaptation aftereffects for some high-level contents provides good prima facie evidence that such contents are perceptually represented.

II.6. Limits To Common Hallmark Arguments

There are two potential dangers for common hallmark arguments. These are not fatal, but point to ways the argument could be sharpened, and to what is still missing for a decisive argument in favor of perceptual Liberalism.

II.6.1. Pitfall #1: Not Distinctive Enough Hallmarks

The first pitfall to be avoided is that the hallmark chosen is not distinctive enough of perception, because it would be shared with some cognitive processes, a pitfall Madeleine Ransom identifies as the "converse fallacy" (2019). Simply, if a hallmark is also exhibited by a cognitive process, then *it* is not a proper hallmark of perception. We can be slightly more subtle though: it might still be a satisfactory hallmark if it is statistically more probable to be exhibited in perception than in cognition (i.e., there are relatively more perceptual states that exhibit it than cognitive states). The converse fallacy is thus more problematic when *many* cognitive states exhibit the same hallmark, and when *few* perceptual states exhibit it.

Intuitions and associations can be very fast and be phenomenologically experienced as instantaneous. Post-perceptual judgments might be fine-grainedly stimulus-dependent just like perception. Some thoughts are said to be *intrusive* in that they occur automatically, without being controllable and dismissible by the subject. Some authors have pointed out that adaptation aftereffects are also displayed in cognition (Helton, 2016; Storrs, 2015). This can be intuited quite easily: if you have been living in New York all your life, then London housing prices won't impress you that much (a kind of *cognitive* desensitization). Contrarily, if you are presented with a house in Kansas City of an average price, you will probably find it quite cheap (a *cognitive* repulsive aftereffect). Closer to our concerns, one can distinguish between adaptation of perceptual representations themselves, and adaptation of *decision criteria* that are used to interpret such perceptual representations (i.e., post-perceptual judgments). Thus, adaptation aftereffects observed for facial expressions could be explained by such changes in decisional criteria (Storrs, 2015): after having post-perceptually judged that 'the FACE EXPRESSES ANGER', the decisional threshold for such judgment is increased, thus lowering the probability of such judgement (everything else being equal). If one is thereafter presented with a *neutral* face, then because of this heightened decision threshold for attribution of anger, one will tend to attribute the opposite expression ('FEAR').

There could be two response strategies for a Liberal to counter such converse fallacy. First, she might try to *refine* the hallmarks that are used. Speed, stimulus-dependence,

automaticity or adaptation are all psychologically intuitive, but they are not very precise and might thus easily fall prey to the converse fallacy. One could thus aim at decomposing such psychological concepts into more fine-grained mechanisms.

For instance, Joulia Smortchkova (2021) has proposed that cognitive and perceptual adaptation effects can be behaviorally distinguished, and rely on quite different psychological processes. In particular, she points out that cognitive "adaptation" can be understood as a classical phenomenon of "contrast effects" (the idea that we tend to judge objects and people relatively to objects and people that are similar, instead of judging them individually), and particularly of "anchor biases" (the idea we tend to judge relatively to the *first* piece of information we receive about an object or people). Interestingly, it is well-known from the anchor bias literature that one can be *de-biased* by several top-down methods (e.g., considering information going in the opposite direction than the anchor, or being alerted about our own anchoring biases). Such de-biasing is *not* observed in perceptual adaptation: if one is adapted to a red patch, one will inevitably see a white patch as green (but will not necessarily *judge* that 'it is GREEN'). No de-biasing of perception is possible by convincing oneself that we are biased. This recalcitrance to de-biasing is linked to the *automaticity* hallmark, and maybe (but much more contentiously, as I have argued) to the informationally encapsulated and impenetrable nature of perception.

This kind of decomposition strategy is theoretically applicable to other hallmarks as well. Both cognition and perception might be indistinguishably fast at a *behavioral* level, but perception might be systematically *faster* at a neural level, and so research that considers *neural speed* might be well placed to use this hallmark. Perceptual stimulus-dependence might be more fine-grained than post-perceptual stimulus-dependence. For instance, perceptual stimulus-dependence might follow laws of psychophysics (e.g., Weber-Fechner law) that it is improbable post-perceptual judgements do follow as well (Beck, 2019; Block, 2023). Finally, we can reasonably suspect, like for adaptation, that cognitive automaticity is amenable to some de-automatization, while perceptual automaticity is clearly not. People suffering from intrusive thoughts, say because of obsessional-compulsive disorders or post-traumatic syndromes, can aim through psychiatric therapy to get control over and dismiss such crippling thoughts. The mere fact that such therapies sometimes succeed is good evidence for the idea that cognitive automaticity is different from the perceptual one.

The second response strategy against the converse fallacy is to point out that while hallmarks might not be distinctive individually, they still are *conjunctively*. Hallmark H_1, H_2, H_3, might thus all be shared by perception and cognition, but it might be that the *conjunction* H_1 & H_2 & H_3 is found only in perception. Fast, stimulus-dependent, automatic, and adapting representations might thus be only found in the realm of perception. One could thus devise experiments that test for several of these hallmarks at the same time and for the same high-level content. While some experiments have this multi-hallmark character (e.g., Li et al., 2002, display both the speed and the parallel processing hallmarks), much more experimental work needs to be done in this direction. As things stand, most common hallmark arguments utilize unique hallmark and high-level property pairs. Speed has been

used for kind contents like naturalness, animal, vehicle; stimulus-dependence for social contents like animacy or relational properties like causation; automaticity for high-level representations of "experts" like Greebles, fingerprints or faces; and adaptation aftereffect for emotional expressions. Adopting a more conjunctive approach (showing that one high-level content can simultaneously exhibit several hallmarks) would be a powerful step forward for common markers arguments in favor of Liberalism.

Again, it should even be expected that such hallmarks be conjoined in some systematic way, if one accepts that perception be defined as some kind of "homeostatic property cluster" (Boyd, 1991), i.e., an ensemble of co-occurring processes that usually (but not necessarily) co-occur because of shared underlying causal factors (Taylor, 2020). Common hallmark arguments would thus be strengthened if we could get clearer on the causal mechanisms that govern the existence of such (conjoined) hallmarks. We would then not merely have an abductive argument where invoking perceptual Liberalism is warranted because it explains some empirical observations, but we would also get a genuine explanation of *why* such perceptual hallmarks are expected to cluster together if perceptual Liberalism is true. Elucidating what these underlying causal mechanisms are will be the central topic of Chapter IV).

II.6.2. Pitfall #2: Low-Level Proxies

The second pitfall we need to be careful about is that we might observe perceptual hallmarks in tasks purportedly requiring representing high-level representations not because such representations really are perceptual, but because such tasks can be "hacked" so to speak and can be completed by representing *low-level proxies* of such high-level contents – what Ransom also identifies as the "low-level confound fallacy" (2019). The existence of such low-level confounds or proxies for high-level contents is ensured by the fact that high-level contents are dependent on the representation of some low-level contents. A Conservative could then argue that what subjects are doing in experiments requiring recognizing high-level contents is to perceptually detect their low-level dependence basis. It would then just be a trivial and uninteresting finding that such experiments exhibit perceptual hallmarks.

Those handling common marker arguments are aware of this issue, and we saw that they can devise clever techniques to try to keep low-level contents constant while making high-level content change. If the studied hallmark is diminished or disappears when the high-level content is taken out, this is a good indication that it was really the high-level content that was driving the hallmark, and not its low-level proxies. This is exactly what Butler and collaborators did (Butler et al., 2008) when they scrambled images of faces, keeping their low-level contents intact but making unrecognizable facial expressions in them. They showed that adaptation aftereffects tended to greatly diminish in such configurations. The adaptation hallmark was thus most probably driven by the high-level facial expression content.

Now, the issue is that we might fear that these precautions will not be deemed sufficient by the Conservative, because of the tremendous amount of low-level confounds that should be controlled for the argument to go through. Worse, it might well be that the low-level proxy is such a complex property that it is impossible to devise it, and thus to control it, a priori. Take the case of adaptation to facial expressions again. While scrambling, rotating, or hybridizing pictures of faces preserves *some* low-level contents, it is also quite clear that these procedures alter or destroy some others. One can for instance think of spatial organizational arrangements of parts, cardinal orientation, or "mid-level" Gestalt contents such as symmetry, which are clearly altered by the scrambling, rotations, or hybridization procedures. The Conservative could here complain that the disappearance of adaptation aftereffects during task completion is merely due to the loss of such low- or "mid"-level proxies, *not* to the loss of high-level contents (Burge, 2014).

This puzzle applies to other hallmarks as well. When subjects recognize scenes as natural or artificial at ultra-fast speed (VanRullen & Thorpe, 2001), it might be because they are able to detect some low-level proxy for such contents, such as the shape and spectral information in a scene, as proposed precisely in Aude Oliva and Antonio Torralba's "spatial envelope" theory of fast scene recognition (Oliva & Torralba, 2001).

In the studies of fine-grained stimulus-dependence presented by Albert Michotte (for causality) or Brian Scholl et al. (for animacy), low-level proxies are controlled for merely by the intuition that situations in which subjects do or do not represent the target high-level content have *very similar* low-level contents. This is the whole point of the "fine-grained dependence" after all. Remember that in Michotte's display, we only vary very slightly an interval delay between motion events to go from subjects representing causation to not representing it. In Scholl's display, we only vary very slightly the motion pattern of the pursuing wolf stimulus to go from subjects representing animacy and chasing to not representing them. But still, there *are* some low-level contents that are varied, and it will always be a lively possibility for the Conservative to claim that successful recognition in these tasks is very finely tuned to some *low-level* changes, so that tweaking with them can dramatically alter task performance, without having to postulate the perceptual representation of high-level contents.

Finally, the automatic representation of some high-level contents (e.g., face) might just be revealing the automatic representation of some complex set of low-level contents that are proxies for such high-level contents. The argument here is similar to the one for adaptation to facial expressions: studies that find a loss of the automaticity hallmark when pictures of the target high-level contents are scrambled, rotated or hybridized (Stokes, 2021a) might be interpreted as merely showing that such transformations have altered some relevant low-level proxies for such high-level contents.

Determining which low-level confounds should be controlled for, and how we should do so, should be a matter of empirical importance for liberal psychologists who wish to use a common hallmark strategy, and we can hope here that scientific progress will allow us to get clearer on this matter.

Nonetheless, we can legitimately worry whether such progress will *ever* satisfy the Conservative. In fact, it seems that for any new low-level proxy that is controlled for, it will be possible to devise a new one that is as arbitrarily complex and fine-grained as a Conservative can conceive of and complain that such proxy has yet not been controlled for. Again, this possibility will be permanently lively for the Conservative because of the articulateness of high-level representations with low-level ones.

II.7. In Search of a Psychological Mechanism

These issues call for some pondering over the common hallmark strategy. So long as the common hallmark strategy remains solely an abduction to the best explanation, I suspect that Conservatives will systematically push for alternatives (be it through appeals to the low-level confound fallacy, the converse fallacy, or a combination of both). A Conservative could always appeal to Occam's razor: it is more parsimonious to clearly separate between the low-level representational capacities of perception from the high-level representational capacities of cognition, and all data advertised by Liberals are explainable within this Conservative framework. Liberals on the contrary introduce a puzzling and non-parsimonious representational capacity in perception in a move that might appear overly *ad hoc*, since high-level content is only postulated to account for common hallmark observations.

The Conservative thus keeps some edge here, and as things stand in the literature, it seems hopeless that common hallmark arguments will come close to being sufficient to arrive at a resolution of the admissible contents of perception debate. We need to go beyond common hallmark arguments' abductive nature and related suspicions of *ad hoc-ness*. We need a *psychological theory.* We need to formulate a proper psychological mechanism that would explain how perception can come to have a representational capacity to represent high-level contents, and why such capacities yield such behavioral hallmarks. If we can come up with such a theory, then the Conservative's Occam razor should lose much of its sharpness. If we have a psychological theory as to how perception can have the capacity to represent high-level contents, then clearly the simplest explanation for common hallmark findings is that such capacity is at work (barring we reasonably keep in check the low-level confound and the converse fallacies). The burden of proof would be shifted into the Conservative camp.

This additional theoretical move in favor of Liberalism should be viewed as a necessary complement to the common hallmark strategy. Taken together, they form a virtuous argumentative couple. Common marker arguments give us good prima facie empirical reasons to believe that some high-level contents are perceptual. Psychological theory explains how perception achieves such representational capacity, and why it displays such hallmarks. Eventually, such psychological theorizing further warrants the empirical

search for these common hallmarks. The next two chapters are devoted to building such a psychological theory for high-level perception.

III. Classification, Recognition, Categorization

All the philosophical agitation surrounding the issue of the admissible contents of perception might probably raise some perplexed eyebrows within psychology laboratories. Most contemporary introductory textbooks on perceptual psychology concur that perception would be rather adaptively worthless if it didn't provide organisms with some meaningful high-level information. Here is for instance Stephen Palmer in his seminal introductory textbook in visual psychology:

> Our perceptual construction of the external world goes even further than completing unseen surfaces in a three dimensional model, however. They include information about the meaning or functional significance of objects and situations. We perceive an object not just as having a particular shape and being in a particular location, but as a person, a dog, a house, or whatever. Being able to classify (or recognize or identify) objects as members of known categories allows us to respond to them in appropriate ways because it gives us access to vast amounts of information that we have stored from previous experiences with similar objects. (Palmer, 1999, p. 13)

There now exists a whole section of perceptual psychology and neurosciences dedicated to the study of the neural and computational underpinnings of how we can perceptually recognize objects, a research program diversely known as "visual categorization" (Serre, 2016), "visual cognition" (Cavanagh, 2011), "perceptual object understanding" (Gauthier & Tarr, 2016), "object recognition" (DiCarlo et al., 2012; Riesenhuber & Poggio, 2002), or "high-level vision" (Cox, 2014).

Most generally, perceptual recognition is scientifically understood as a two-staged process: first, perception extracts the fine-grained low-level information contained in the stimulus (what is known as "early vision"); then, low-level information is *interpreted* by various matching algorithms which associate it with categorical labels. Debates then arise as to which matching algorithm is used, and how to computationally describe it. The scientific literature usually takes visual recognition, identification, categorization or understanding as more or less equivalently designating such high-level interpretative mechanism. As I will argue in this chapter, I think this threatens to equivocate several different mechanisms that are functionally and representationally different in nature.

Perceptual Liberalism is thus widespread if not trivial within the scientific community. If so, then the requirement we ended up with in the last chapter finds a direct solution: to get the psychological mechanism that underwrites high-level representations in vision (and that explains common hallmarks observations), we simply can look for the currently best psychological models for visual categorization. At this point, one might even wonder why the admissible contents of perception debate still survives in the philosophical community, since it appears that Liberalism has assuredly won in the lab.

Matters are not so straightforward though. Talks of visual categorization as some kind of *interpretative* process should make us cautious: do scientists appropriately distinguish between genuinely perceptual representations and their *post-perceptual* interpretations (post-perceptual judgments)? We have good reasons to be suspicious, especially considering that the explanatory goals of scientists and philosophers are quite divergent on this matter. While building a clear-cut distinction between perceptual and cognitive representations (seeing-*as* and seeing-*that*) bears many philosophical implications, it is less clear that such a distinction is so much relevant if one aims at explaining human behavior in general.

It shouldn't be surprising then that psychologists easily accept a view where "high-level vision" is understood as a hybrid representational state standing in-between early vision and high-level cognition (explicit deliberative thinking), as exposed by such exotic denominations as "visual cognition". Some psychologists even explicitly deny that there is anything like a discrete distinction between perception and cognition, both standing rather on a mental continuum with no clear borders (Goldstone & Barsalou, 1998; Quinn & Eimas, 1997). As things stand then, it is not conspicuous that the high-level vision research program upholds strictly *perceptual* Liberalism. As I argue in this chapter, I think classical mechanisms that psychologists and neuroscientists advance to describe "visual" categorization only describe some kind of representationally hybrid, post-perceptual mechanism.

In this chapter, I will first make some theoretical considerations as to how we should understand the representational border between *perception* and *cognition* in order to have some criterion for deciding whether a psychological is or isn't genuinely *perceptual* (III.1). I then argue that most of the psychological literature on "high-level vision" (and synonyms) actually describes a process of *conceptualization* of perceptual representations, a process for which I reserve the term "categorization". I argue that categorization violates the criteria for perception exposed previously, and thus is not a good model for the psychological mechanism we are looking for to defend Liberalism (III.2). However, the pre-conceptualization or pre-categorization perceptual processes described in the literature will come out as sufficient to ground some kind of nonconceptual *but still high-level* kind of content, vindicating perceptual Liberalism. While the consideration of the psychological literature will first look disappointing for Liberalism, further considerations will demonstrate crucial import from it.

III.1. What Representational Mechanisms should count as Perceptual?

III.1.1. The Non-Conceptuality Criterion

I have stayed away from characterizing what I take to distinguish perception and cognition. This is because clarifying such distinction (i.e., defining necessary and sufficient conditions for perception and cognition) is a highly debated topic (Nes et al., 2023), and favoring one criteria above others in the ambit of the admissible contents of perception debate might too easily lead to question-begging strategies (Conservatives choosing very restrictive criteria for perception while Liberals will choose more encompassing ones).

We saw one can try to get around such definitional controversy by relying more consensually on *behavioral hallmarks* of perception and determining whether these are exhibited by some high-level representations. While this common hallmark strategy gives good prima facie reasons to favor Liberalism, we saw that on its own it will systematically run against plausible objections. To counter such objections efficiently and systematically, Liberals should make explicit a psychological mechanism that grounds the capacity to perceive high-level contents. But now, proposing such a mechanism cannot be done without at some point facing the difficult question of what constitutes a *perceptual* psychological mechanism, as opposed to a *cognitive* one.

The solution I propose to build such distinction without ending up in question-begging quagmires is the following: we can begin the investigation by building some *maximally restrictive and empirically plausible* criteria for perception, at least *maximally restrictive* in the sense that they are acceptable by a Conservative while not making Liberalism evident, and *empirically plausible* in that they do not violate basic commitments in psychological science. If we manage to show that *even under such restrictive criteria*, we can still get representations of high-level contents, then this mechanism should be acceptable, even for the Conservatively-minded.

At no point during the argument will I consider these restrictive criteria as *definitional* of perception. I am *not* claiming that they constitute in any sense jointly necessary and sufficient conditions for a mechanism being perceptual. Rather, these criteria will be *operational*: they should be viewed as criteria that reasonably distinguish between perception and cognition *without begging the question of the admissible contents debate*.

The central criterion I propose to distinguish between perception and cognition in the ambit of the admissible contents debate is a classical one in contemporary philosophy of perception (Block, 2023; Burge, 2022): I propose that we view perception has computing over *nonconceptual* contents, while cognition computes over *conceptual* contents. My advocating for this criterion should not be taken as a stipulation that perception is by nature nonconceptual. Again, my criterion is only operational: I will *suppose* that

perception is nonconceptual, and from there show that we can still get to some high-level representations in perception.

I think this is a relevant criterion in the ambit of this discussion since it a priori restricts the representational capacities of perception to a minimum. If we were to consider that perception could compute over conceptual contents, then Liberalism might seem trivial. I take concepts to be mental representations. As representations, concepts can clearly have high-level contents. Typical examples of concepts include CAT, WATER, CAR or JUSTICE. As we saw in I.2 SEEING AS SEEING-AS, the doxastic view of perception maintains that we can see something as a CAT, WATER, or a CAR, because we can see *that* such thing 'is a CAT', 'is WATER', or 'is a CAR'. Such doxastic view of perception implies that perception is conceptual in nature since beliefs are understood as compositions over concepts. Doxastic views of perception trivialize Liberalism, because our beliefs about what we see can clearly contain high-level conceptual contents.

One could be a conceptualist about perception without upholding such doxastic view. In particular, one could think that perception can token *purely demonstrative* concepts (McDowell, 1994). When we see something as red, maybe this is equivalent to perceptually tokening the concept 'COLORED THUS' (with "thus" expressing a demonstrative perceptual relation to the color). In another formalism, one could be said to see something as THIS(color). Once one loses the demonstrative perceptual relation to the color (e.g., when one closes her eyes or when the colored object disappears), the concept is not tokened anymore. Such purely demonstrative concepts do not seem to trivialize Liberalism. In fact, a Conservative could accept that perception possesses such conceptual representations. Conceptually seeing things as THIS(color), THIS(shape), or THIS(size), where the "this" demonstrative relation is realized perceptually, is compatible with the general idea that perception only represents low-level contents. In that case, the contents would be low-level demonstrative concepts.

If purely demonstrative concepts could be accepted in perception even by a Conservative, why should we take *nonconceptual* contents as the relevant criterion to demarcate perception from cognition?

My main answer is that I think that such *purely demonstrative* understanding of concepts is overly revisionary and unintuitive. Concept possession is supposed to be some kind of high intellectual achievement. Animals and infants are supposed to not be concept possessors (or at least, to possess a very limited set of concepts). Accepting a purely demonstrative view of concepts goes against such understanding. In fact, the condition on concept possession is (too) much relaxed, as it only requires a perceptual relation to an object and its properties as well as the possession of a re-applicable demonstrative representational capacity. Such requirements are potentially fulfilled even by animals and infants.

Furthermore, such a view theoretically goes against the grain of some basic commitments in how most linguists and psychologists understand the idea of "concept". Such commitments are extremely varied, but among them we can count *the capacity to*

categorize objects and to re-identify and re-apply the same concept to different particulars, *the capacity to draw inferences*, and *the possibility to think with concepts in the total absence of the referred particulars and properties, making possible communicating about such particulars in their absence*. Psychologist Gregory L. Murphy for instance introduces his seminal review of the psychology of concepts by claiming:

> The psychology of concepts, then, has the goal of understanding the representations that allow us to do all these things, most importantly, identifying objects and events as being in a certain category, drawing inferences about novel entities, and communicating about them. (G. L. Murphy, 2004, p. 3)

Purely demonstrative concepts seem to tamper with such commitments. They thus can be viewed as advancing an overly *ad hoc* and revisionary understanding of what "concepts" are.

Conceptualists about perception could be tempted here to argue that demonstrative low-level concepts like THIS(color) actually do fulfil these criteria. They yield the possibility of re-identification and re-application of the concept. Once I have applied the concept THIS(color), I might be able to re-apply it to another shade of color and recognize as the same THIS(color). This view though is empirically dubious to say the least, as Sean Dorrance Kelly has rightly pointed out (Kelly, 2001). It is both theoretically conceivable and empirically hardly doubtable that it is possible for one to perceive a certain shade of color but be incapable of re-identifying it as that same shade afterward. One could see a color as THIS(red) but fail to see the same shade as THIS(red) the next day, or fail to think about the same THIS(red) in the absence of the color because one simply does not remember the exact shade that was perceived. One can thus doubt that demonstrative concepts can really obey the re-identification and distance-from-referent criterion. As Kelly I think rightly concludes:

> [It] looks as though we can make (or at least we can imagine making) discriminations in perception, and hence can entertain distinct perceptual contents, in cases in which we don't satisfy the [re-identification] criteria for possessing the demonstrative concepts for the things discriminated. Demonstrative concepts, then, cannot in general characterize the content of perceptual experience. (Kelly, 2001, p. 412-413)

Finally, and most importantly, I simply reject the demonstrative understanding of concepts because I think this understanding is not operational in the ambit of the admissible contents debate. In fact, *what* properties we can perceptually be demonstratively related to *is precisely the question we are trying to answer*. Saying that we can have purely demonstrative concepts in perception tells us nothing about whether we can only have low-level demonstrative concepts like THIS(red), or whether we can also have high-level demonstrative concepts like THIS(leopard). Answering this question simply amounts to answering the admissible contents of perception debate. I'm thus not certain that the debate over the possibility of

demonstrative concepts in thought is relevant to the problem at hand. I thus exclude the idea of demonstrative concepts as constitutive of perception in what's to follow.

The nonconceptuality criterion is in any case not primarily about demonstrative concepts. Rather, it is about non-demonstrative, *general* concepts like LEOPARDS, CARS, or JOE BIDEN. I want to argue that perception should be taken as not being representationally able to have such general conceptual contents.

What exactly are those general non-demonstrative conceptual contents? And oppositely, what are *nonconceptual* contents? In general, we can view general concepts as the atomic representational constituents of propositions. Propositional representations are representations with a certain structure (or syntax) that behaves in a language-like way. Such syntax is thought to be able to implement basic logical operations such as subject-attribute predication, logical connection (negation, disjunction, conjunction, implication, equivalence), and quantification. Bound to this idea, one can think of concepts as standing to propositions like words stand to sentences. This idea is nicely put by Christopher Gauker:

> For each judgment there will be a sentence that can be used to make an assertion that expresses it, and that judgment will have components that stand to the whole judgment as the open class words of the sentence that can be used to express it stand to the whole sentence. Concepts are *such* components of judgments. In saying this, we do not *define* concepts in terms of language, but we use a certain conception of language to locate the subject matter of a theory of concepts. (Gauker, 2011, p. 3)

(Here, we could innocuously replace "judgment" with "proposition").

Now, we can consider what kind of constituents usually form part of propositions. Here are some typical samples of statements: "My dog is black", "The key is on the nightstand", "Quentin is tired". These statements are informative because they relate to different *bodies of information*. Compare with non-typical samples of statements: "This dog is this black", "This key is on this nightstand". These statements have demonstrative constituents: what they express depends on the synchronous demonstrative relation of the locutor with some worldly objects or properties, in the presented case: a perceptual relation. These non-general statements are composed of demonstrative concepts.

One should think of general concepts as constituents like in the former, non-demonstrative statements. Concepts retrieve bodies of information that are put together and stored to be recovered at any time to build general propositions. Crucially, they can be recovered *offline*. I can think the proposition "the key is on the nightstand" without being right now related to the key or the nightstand. I am here following on the view of propositions, and thus of conceptual contents, developed by Tyler Burge, who puts the matter clearly:

> What is important and distinctive about attributives in propositional structures—*conceptual attributives*—is (a) that every one of them *can* occur in a representational content outside the scope of any referential application, and (b) that in every

> propositional structure at least one of them does occur outside such scope. (Burge, 2022, p. 205, fn. 199)

Conceptual contents are those representational contents that can be tokened in the absence of a demonstrative referential relation, such as, crucially, the demonstrative referential relation provided by sense perception. As Burge put it, "[p]erception does not involve a competence to characterize that is independent of serving perceptual reference" (Burge, 2022, p. 202), whereas (non-demonstrative) thought involving concepts *can* fundamentally function independently of demonstrative referential functions. One can think that 'CATS are FURRY' without at the same time demonstratively referring to particulars (*cats*) or properties (*furriness*). This thought is semantically meaningful (obeys accuracy conditions) because the concepts CATS and FURRY retrieve a body of information that is stored by the individual (or by experts in the individual's linguistic context, cf. Burge, 1979), which allow to get at some referred to particulars and properties. As Edouard Machery put it:

> A concept of x is a body of knowledge about x that is stored in long-term memory and that is used by default in the processes underlying most, if not all, higher cognitive competences when these processes result in judgments about x. (Machery, 2009, p. 12)

It is the fact that general concepts are viewed as such offline bodies of knowledge (or, more accurately, as flexibly *retrieving* or *pointing* to such bodies of knowledge) that make them relevant for higher cognitive competences like thought and reasoning. Thinking can be viewed as organizing such bodies of knowledge (i.e., drawing links between them, updating them, erasing them). We wouldn't think to the full human potential if we could only think through demonstrative concepts.

In close connection with the preceding points, we can point out that the general concepts constitutive of propositions can be tokened in an amodal way[38]. One can think that 'QUENTIN is TIRED' without retrieving any modally-specific (visual, auditory, olfactory or what not) information about QUENTIN or BEING TIRED. It suffices that one points to some amodal bodies of information linked to the concepts QUENTIN and BEING TIRED. No more dramatical illustration of this idea could be given other than the fact that blind subjects can master color concepts (J. S. Kim et al., 2021; Saysani et al., 2021). Blind patients can think that 'FIRE HYDRANTS are RED', though they store no modally-specific information about fire hydrants or red.

Now, general concepts *can* be linked to modally-specific information. The general concept DOG is linked to information both in modally-specific formats (e.g., the sound of dog barks), but also information in amodal format (e.g., the information that dogs are mammals, that they are friendly, that they are genetically related to wolves). Blind people

[38] I think this is a reasonable view of thought, though some philosophers and psychologists have argued that *all* conceptual tokenings are grounded in modally-specific sensory simulations (Barsalou, 2010; Prinz, 2002). I won't dwell on this view, which is both philosophically highly contentious and empirically dubious (Shallice & Cooper, 2013).

might store enough amodal information about dogs so that they can be said to possess the DOG concept, even though they lack some modal information about them (J. S. Kim et al., 2019). The point is thus that general concepts can point to *both* amodal and modal information.

Oppositely, I think that there are concepts that might be purely *observational*, when the only information they are linked to is modally-specific information. One might thus have a purely observational concept of RED. One shouldn't confuse the *pointer* and the *information it points to* though. Observational concepts are still concepts in that they can be tokened non-demonstratively in propositional structures. One can think that 'CATS are CUTE' by retrieving only perceptual information about cats. One can think so with ones' eyes closed. Even when observational concepts are predicated *demonstratively*, as when one thinks 'this is a CAT', we are still dealing with some kind of (proto)proposition relying on some (proto)predication. This demonstrative thought might not be a full proposition with a full predicative composition because it predicates a concept not to another concept, but directly to a demonstrated object (an object demonstrated anaphorically through perceptual referential capacities). It *still* uses a conceptual capacity though (retrieving and predicating the observational concept CAT), and thus on my view cannot be considered as a *perceptual* representation[39].

Furthermore, the observational concept RED is not reducible to the information it points to. It might come that the same observational concept RED might be at some point linked to amodal information (e.g., that 'RED can symbolize DANGER, PASSION, and COURAGE'), thus becoming a non-observational concept. Still, it is the same concept, the same amodal semantic pointer.

Perception does not recover its representational constituents from offline, memorized bodies of information. Perception is a demonstrative capacity that puts us in direct contact with its referents. Perceptual representations have a demonstrative nature in that what they represent depends on the synchronic relation of the subject to the represented particular at the moment of the tokening of the representation[40].

Perception does not compute over non-demonstrative, amodal general concepts. Perception is demonstrative and modally specific. Any perceptual representation depends on a specific sensory modality[41]. We should take perception as being nonconceptual (both rejecting *demonstrative* and *general* conceptual contents). In our search for a convincing psychological mechanism that could uphold perceptual Liberalism, I will thus only consider psychological mechanisms that employ nonconceptual contents. If Liberalism is to be

[39] Notice that there a difference between the *demonstrative perceptual concept* THIS(cat), a kind of representation I reject, and the *demonstrative thought* 'this is a CAT', which I think are not perceptual, but hybrid perceptual-cognitive representations relying on observational conceptual predication.

[40] One can surely *imagine* the Eiffel Tower in an offline way, but in this thesis, I am following consensus in setting apart imagination and perception.

[41] The existence of *cross*-modal representations (e.g., audiovisual representations) is compatible with this claim, since cross-modality is not amodality. This is compatible with the nonconceptual nature of perception.

vindicated, it will not be because we can see something as a LEOPARD (that would break the nonconceptuality criterion), but because we can see something as a leopard (remains to be determined what exactly is the semantic nature of such high-level *nonconceptual* content).

III.1.2. Semantic and Perceptual Memory

These latter considerations about the nonconceptual nature of perception point to an important idea: *perception does not include access to semantic knowledge as stored in semantic memory*. I take "semantic stores", "semantic memory", or "semantic hubs" to be representational systems linking memory representations *in a transmodal way*. Storing representations in a transmodal way means that one stores linked information about something in both modally-specific ways *and* in amodal (mostly, verbal) ways. These memory representations linked to transmodal information are what I call *concepts*. Thinking is tokening concepts and thus retrieving transmodal information. So, for example, the concept CAT is stored in semantic memory, and is linked to some perceptual information (visual, auditory, tactile) but also to amodal information (that 'CATS are ANIMALS', that 'CATS LIKE MILK', that 'CATS are CUTE'). What information a representation in semantic memory is bound to can be updated through learning.

Concepts in semantic memory can be understood as "semantic pointers". They are some kinds of unstructured representations which are associated with different kinds of information in modal or amodal format. These semantic pointers are themselves amodal (their tokening is not constituted by a specific modality, but it is purely cognitive), but they can point to modally-specific information (e.g., what cats look like) as stored in perceptual forms of memory. Jake Quilty-Dunn, updating the atomic theory of Fodor (Fodor, 1998), has recently developed such a "pointer" view of concepts, claiming that "concepts are syntactically atomic representations that address memory locations. At those memory locations are stored a large array of representations, including prototypes, theories, and other informational structures" (Quilty-Dunn, 2021, p. 173). This view is backed up by computational models in cognitive science (Blouw et al., 2016) and neuroscientific findings demonstrating transmodal activity in specific areas (especially the anterior temporal lobe) in concept-hungry tasks (Ralph et al., 2017). Crucially, representations in semantic memory can be tokened in the absence of represented particulars and properties. One can think 'CATS are CUTE' in the absence of any denoted cats or cuteness. Semantic pointers can also retrieve information in a context- and task-relative way. I don't think that concepts necessarily retrieve modally-specific information, even though they can.

When I say that perception is nonconceptual, I thus claim that perception does not have access in its computations to such semantic pointers and their associated transmodal information. Perception remains modally-specific from beginning to end. Of course, perception must be connected to semantic stores in some way, at least in a bottom-up way (and probably also in a top-down way). When we perceive something, we can reason about

what we see and bound our current perceptual representation to specific concepts stored in semantic memory. Perception is doubtless a primary way for updating our semantic stores. I can also think and reason about what I'm currently seeing by retrieving information from semantic memory. When I see a dog, I can retrieve from semantic memory my own concept of DOG, and information attached to it, e.g., that 'DOG can be DANGEROUS', the auditory information as of bark sound etc. Such transmodal activation is typical of semantic cognition. It takes us out of the realm of modally-specific, nonconceptual perception. Any theory of high-level perception that appeals to semantic stores thus should be viewed with cautious as describing at best some *impure* or *hybrid* perceptual-cognitive process.

The isolation of "pure" perceptual processes from semantic stores does *not* imply that perception only representationally works with its synchronic sensory inputs. Perception has access to proprietary memory traces to perform its representational functions. But memory stores implied in perception are to be strictly distinguished from semantic stores of conceptual pointers. Perception hosts proprietary forms of memory (Luck & Hollingworth, 2008; Schurgin, 2018). Perceptual forms of memory (including iconic memory, visual working memory and visual long-term memory) store *perceptual* memory traces. Perceptual memory traces, like visual perceptions, have nonconceptual contents, while semantic memory stores transmodal concepts.

I think long-term perceptual memory is particularly important for understanding high-level perception. A good illustration of long-term perceptual memory is displayed in studies where subjects must memorize in the order of thousands of pictures to complete a recognition task. In such studies, subjects are able to determine whether a test picture was previously presented or is new even if they had to retain thousands of images and even though the test is done days after the training session (Brady et al., 2008, 2011; Schurgin, 2018). Such result suggests a massive visual memory storage capacity. Furthermore, it was shown that such storage is achieved with a very high level of perceptual detail or "fidelity". Subjects can determine which of two images was previously seen and which is new, *even if* the only difference between the two images is a subtle visual detail (e.g., position, color, or rotation of objects).

Such high-storage, high-fidelity capacity puts long-term perceptual memory aside from the more classical ideas of both semantic and episodic memory. Contrary to semantic memory, visual memory stores modally-specific, high-fidelity visual information. It does not streamline information through propositional-conceptual contents ('MARY LOVES CATS' 'MARY LIVES IN NAPLES'), but rather preserves visual nonconceptual traces of what was seen (i.e., at least shape, color, motions… for a Conservative, and high-level *nonconceptual* contents for a Liberal).

Contrary to *episodic* memory though (but like semantic memory), perceptual "memory" is not a representational capacity for recollecting one's own spatiotemporally specific past experiences. Perceptual memory traces are not bound to time-specific events in the past, and nor bound to one's own perspective on them. When one retrieves a

perceptual memory (e.g., retrieves what Tasmanian devils look like), the retrieved perceptual representation is not indexed to a specific timepoint and to a specific egocentric viewing perspective. Rather, one retrieves a perceptual representation that is temporally neutral and is non-perspectival. James Openshaw has well described this perspectiveless nature of what he calls "object remembering":

> [O]bjectual recall may more closely resemble semantic memory than episodic memory. But objectual recall involves activating a certain kind of (singular) knowledge in a distinctive way. It is constitutively bound up with the subject's retention of abilities to engage in conscious imagery which appropriately matches the object's past perceptible qualities, generalised from across the times at which the subject encountered the object. (Openshaw, 2022, p. 11-12)

Perceptual memory and associated "object recall" abilities as described by Openshaw are thus quite unlike the traditional view of memory as perspectival representations of the subject's past and might be more adequately understood as a kind of perceptual "knowledge" that shares the perspectiveless character of semantic knowledge. I will keep with the term "perceptual memory" though as this is a common practice in the scientific literature (Luck & Hollingworth, 2008).

In the case of subjects recognizing pictures as previously seen or unseen as described above, we can conjecture that their high success in this task relies on some perceptual memory system as distinguished from semantic or episodic memory. Subjects cannot be expected to episodically remember each episode of seeing each individual picture. Nor can they be expected to store thousands of demonstrative concepts corresponding to each picture in semantic memory. When they recognize having already seen a presented picture, this must be because the presented picture matches a perceptual memory trace, and because such matching is accessed in some way (for instance, through a metacognitive feeling of familiarity) by the subjects. But subjects do not have to recall a moment in the past when they perceived the picture (though of course, they could come to recall it), nor do they need to judge that they saw THIS(picture) before.

These points are important because they show that the nonconceptuality criterion is compatible with perception having access to its own proprietary memory (or "knowledge") stores, as distinguished from classical semantic, conceptual stores. This idea will be at the core of our quest for a purely perceptual psychological mechanism for grounding high-level contents, where it will be especially important to distinguish between access to *perceptual* memory from access to *semantic* memory.

Of course, this idea of independent *perceptual* memory stores is perfectly compatible with the idea that memory systems remain strongly interconnected. Semantic memory might store information that is indexed to perceptual memories. For instance, one might know that 'FIRE HYDRANTS are COMPULSORY in PUBLIC SPACES', and conjointly be able to summon from perceptual memory some object memory as of a fire hydrant, thus enriching one's amodal (verbal) information with modally-specific perceptual information about fire

hydrants. Episodic memory is also quite obviously tied to perceptual memory. When one remembers through conscious mental imagery the 10th anniversary of her brother, one might "reconstructively" picture what her brother looked like by using perceptual memory traces. Notice that, crucially, the reconstructed perceptual image of the brother need not necessarily correspond to what one's brother looked like exactly during this party (this is why episodic memory is now often considered to be "constructive" and not "preservative", cf. e.g., De Brigard, 2014). Of course, to count as genuine *episodic* memory, such constructive capacity helped by perceptual memory must be attached to some indexing to some particular time in the subject's own past. Whether such capacity requires some minimal concept possession (e.g., of the egocentric I or the temporal THEN), or whether such capacity might stem from some autonoetic feeling can remain an open issue here. What is important is only the distinction of perceptual memory (which is nonconceptual, not time indexed, and not about oneself) from semantic memory (which is also neither time-indexed nor about oneself, but conceptual) and episodic memory (which is time-indexed and about oneself, and which might or might not be conceptual).

With this nonconceptuality criterion in hand for perception (including perceptual memory), we can now embark on the quest to find a psychological mechanism to ground *perceptual* high-level contents. Are there psychological mechanisms that are at the same time purely reliant on perceptual nonconceptual contents, but that can still ground some kind of *high-level* content? The realm of "high-level vision" or "categorization" seems to be where to look. However, I show in the next section that mechanisms described in the "perceptual categorization" literature most often violate the nonconceptuality criterion, and imply access to semantic stores.

III.2. Scientific Models of "Perceptual" Categorization as Hybrid Perceptual-Cognitive Models

The most natural place to look for a perceptual mechanism grounding high-level contents is the vast empirical literature studying high-level perception. The issue though, as we will see right away, is that we can reasonably distrust the idea that this literature undoubtedly targets *perceptual* mechanisms. Why are scientists so prone to talk about high-level *perception* then (or *perceptual* object recognition, *perceptual* categorization etc.)? I think we should take them as targeting not a specific representational kind, but rather as targeting variegated mental processes operating over sensory inputs. High-level "perception" is an ensemble of processes that *classify* perceptual representations. Such processes can theoretically be perceptual themselves (this is what we are searching for), but they could also be cognitive, or perceptual-cognitive hybrids. For instance, if one is presented with a test image T to be classified – e.g., "is this image an X or a Y?" –, one might simply have some knowledge that 'all things with conjoined features a, b and c are X', and thus think, upon seeing T, *that* 'it is an X' because it has features a, b, and c. In the ambit of this thesis,

such classification should be understood as a post-perceptual judgment, and not as a perceptual representation proper. One sees that 'this T is an X', but one does not necessarily see T as an X.

Since the classified stimulus is an image, the classification task can naturally be viewed as a *perceptual* one, even if the psychological mechanism employed to solve it is not itself perceptual. A good contrast here is the study of *linguistic* classification, or the ability to classify *words* into higher-level categories (Rosch & Mervis, 1975). For instance, one can be tasked with determining whether a word should be classified in the "fish" or "mammal" category. In that case, the input to the classification process is not a perceptual representation, but a linguistic one[42].

The high-level "perception" literature is mostly called such in psychology because it experimentally focuses on the classification of visual inputs, not primarily because it involves purely perceptual mechanisms. One thus cannot straightforwardly take a mechanism in high-level "perception" (e.g., the capacity to classify pictures into animal or tool categories) and conclude that it upholds high-level contents in perception. One should describe some criterion for distinguishing between perception and cognition, and show that the considered mechanism of classification fulfils this criterion. From now on, I will thus consider in some detail various classification mechanisms proposed in the high-level perception literature. I show that most of these mechanisms appeal to conceptual contents, and thus should at best be viewed as hybrid perception-conceptual mechanisms unable to uphold Liberalism. I will specifically reserve the term *categorization* for such hybrid mechanisms.

III.2.1. Classification, Categorization, Recognition

We should note at the outset that in the psychological literature, theories of "*concept* application" and theories of "classification", "recognition" or "categorization" in general are very often equated. This is highly problematic in the context of the admissible content of perception debate though, especially under our understanding that perception should be understood as having *non*conceptual contents. If *any* perceptual classification was to be equated with a kind of concept application, then high-level perceptual content would be made necessarily oxymoronic.

To avoid such confusion, I propose to make from the onset clear terminological distinctions. I reserve the term "classification" for the general subsuming of some particular under some high-level sortal, be it an *individual* sortal or a *class* sortal. I take sortals to be some equivalence class of represented particulars. Sortals characterize particulars as being

[42] Of course, getting to the linguistic representation must go through a perceptual representation at some point (e.g., visually representing written words), but in that case the perceptual representation itself is not thought to be relevant to the classification process (it does not matter how the word is written, what is its color, font, size, etc.).

identical in some ways. All red particulars are identical under the red sortal. All cat particulars are identical under the cat sortal. Subsuming a particular under a sortal (i.e., classifying it), is thus attributing a *kind* property to it.

I reserve "classification" for subsuming of particular under *high-level* sortals to not overextend its range of application. We obviously perceptually subsume particulars under shape sortals, color sortals, gestalt sortals etc. This is not what psychologists or philosophers usually take perceptual "classification", "categorization", or "recognition" to concern. Again, I do not take "high-level" sortals to be definable through clear necessary and sufficient conditions. What count as high-level sortals (cats, cars, Joe Biden, beautiful, graspable…) is taken to be relative to the state of the admissible contents of perception debate (see I.3 WHAT IT TAKES TO BE HIGH IN THE PERCEPTUAL ECONOMY).

Perceptual classification into high-level sortals can be achieved both through *categorization* or through *recognition*. I reserve the term *categorization* for classification through concept application. I reserve the term *recognition* for nonconceptual classifications. **Figure 2** schematizes these important distinctions. I make explicit that classification can be as of individuals or classes of things.

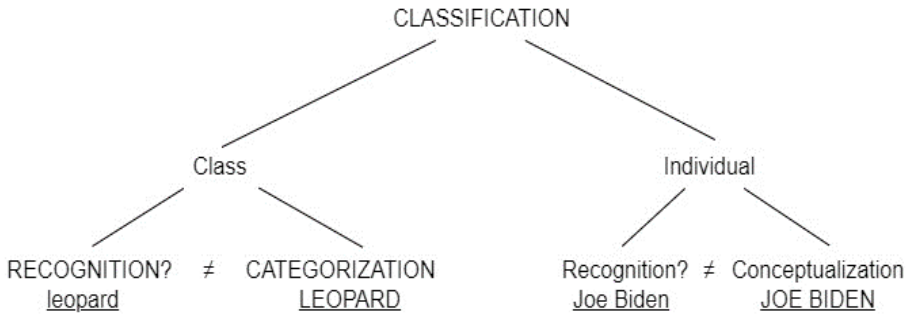

Figure 2. Different classification processes.

I *do not* reserve the term "recognition" merely for the recognition of individuals. One can recognize someone as Joe Biden or recognize something as a leopard (just like one can categorize someone as JOE BIDEN or categorize something as a LEOPARD). I also do not take recognition to necessarily involve a conscious experience (such as a feeling of familiarity). I think recognition *can* yield proprietary kinds of experiences (which are to be distinguished from the cognitive experiences of categorizing something). But one can recognize unconsciously, i.e., unconsciously subsume a particular under a high-level sortal, or equivalently, unconsciously attribute a kind property to it. This terminology does *not* correspond to what scientists call classification, categorization, or recognition. Scientists usually do not differentiate these terms. I think what scientists mostly describe through

these terms is a process of categorization understood as concept-application. The topic of this chapter and the next is to determine whether there exists a psychological mechanism of classification into high-level sortals that is purely *recognitional* in nature (not applying concepts), instead of being *categorical* (applying concepts).

III.2.2. Psychological Models

Many psychological mechanisms have been proposed in the literature to account for perceptual classifications. I will here focus on three central propositions that are by far the most discussed (Machery, 2009; G. L. Murphy, 2004): rule-based *descriptive* mechanisms, similarity-based *prototype* mechanisms, and similarity-based *exemplar* mechanisms.

III.2.2.i. Rule-Based Mechanisms

Rule-based mechanisms of classification designate those psychological mechanisms relying on storing in semantic memory some description of a category. When one is presented with a stimulus to be classified, one checks whether the stimulus fits such stored description or not. A topical illustration of the rule-based mechanism is what psychologist Gregory Murphy refers to as the "classical view of concept" (G. L. Murphy, 2004), or the view that the descriptions used to classify stimuli are *definitions*. Definitions are descriptions stating jointly necessary and sufficient conditions for concept application. For instance, carbon dioxide defined necessarily and sufficiently as a chemical made up of one carbon atom bounded to two oxygen atoms. Anything that fits this definition can be classified as carbon dioxide. While such definitional view of concepts was historically dominant throughout much of pre-20th century psychology and philosophy, it slowly fell out of grace (G. L. Murphy, 2004). Careful philosophical theorizing showed that most concepts are not definable in terms of jointly necessary and sufficient conditions, but rather subsume objects which bear some looser kind of "family resemblance" (Wittgenstein et al., 2009).

Nonetheless, the rule-based descriptive tradition did not disappear. An influential descriptive strategy that is alive today is to understand the descriptions used in classification as being *theories* (Gelman, 2004; Keil, 1992; Mandler, 2000). Roughly, a theory is a description of some properties of objects within a category, attached to some description of the common *hidden essence* of these objects which explains why they share these properties. Such hidden essence can for instance be the common origin of objects, or their common functional goal. For instance, a theory of "vehicles" could have descriptions of some of its superficial features ("it moves", "it can carry people") as well as essential explanations for why it has these features ("vehicles are produced to move people"). Crucially, theories can account for the fuzziness of classification processes since such theories might be themselves vague and allow for limit cases: one's description of features might not determine which of these features are necessary or sufficient for the fulfillment of a certain functional essence. For instance, what features are essential to consider that something is an object made to carry people (a vehicle) can remain vague.

Such "theory views" of classification are supported by results in developmental psychology that demonstrate that children do not stop at superficial features of objects to make classification decisions. Instead, they use presuppositions about how these features are normally bound to some essential property (Gelman, 2003). For instance, if a child learns that a skunk-looking animal was formerly a racoon that has been disguised into a skunk by a playful scientist, or that has physically transformed into a skunk-looking animal because of a strange disease, children classify such animal as a raccoon, even if it superficially looks like a skunk (Keil, 1992). This is interpreted as children possessing a *theory* about animal species, according to which a species is defined by some hidden substance and a particular biological origin and behavior, *not* by their physical appearance.

I won't discuss the merits and issues with such rule-based approaches. What is relevant for us here is that such classification mechanisms are obviously conceptual and thus not perceptual in nature. When one checks whether a currently perceived stimulus satisfies some definition or theory, one is checking whether the currently perceived stimulus fits to a particular *concept* that points to such definition or theory. Rule-based classifications obviously appeal to semantic memory. Neither the application nor the storage of such definitions or theories can be attributed to nonconceptual perception, and rule-based mechanisms clearly describe a cognitive *categorization* (i.e., concept application) process (which they are often explicitly meant to).

III.2.2.ii. Similarity Views

Descriptive mechanisms are traditionally opposed to another family of classification strategies known as *similarity-based* classification (Cheung & Bar, 2014; Edelman, 1999; Goldstone, 1994b; Hahn & Ramscar, 2001; Medin et al., 1993; Nosofsky, 1986; Palmeri, 1997; Rosch & Mervis, 1975; Sloman & Rips, 1998; Sloutsky et al., 2007; E. E. Smith et al., 1998; L. B. Smith & Heise, 1992; Tiedemann et al., 2022). Under such views, classification is achieved not by evaluating whether the current perceptual representation fulfils some description (definition or theory), but by evaluating whether such representation *is similar* to some other representation stored in memory *which classification we already know*. On these views, to classify a target stimulus T, one compares T to stored representations with a specific label, say the representation-label pairs $\{R_1$, "X"$\}$, $\{R_2$, "Y"$\}$, $\{R_3$, "Z"$\}$. Then, one computes a similarity score between T and each such representations: (T-R_1, T-R_2, T-R_3). T inherits the label of the representation that gives the highest similarity score.

To illustrate, one can classify a picture as one of a cat because one sees that this image is maximally similar to some previously seen images labeled as cats (and less similar to previously seen images labeled as dogs, horses, cars, or teacups). Notice that in this case, one does not need to possess any kind of description or theory about cats. This is supposedly one advantage of similarity-based views, since by de-intellectualizing perceptual classification achievements they more naturally account for the fact that we seem to be able to classify an object even when we do not possess any kind of explanation as to *why* we classify it so. We simply classify it by noticing that it is similar to some memorized item.

This achievement is particularly salient when classification is achieved over very artificial classes, such as lab-created classes of random dot patterns for which we probably do not store any kind of specific description (Posner & Keele, 1968). It is also salient in creatures of whom we do not expect that they possess any particularly developed theoretical capacities, such as non-human animals and infants (for a review: J. D. Smith et al., 2016).

Furthermore, it is a *core* prediction of similarity-based views that classification should be fuzzy, and that some stimulus should be deemed as more *typical* of a class than others. Similarity scores are not 0 or 1, and so objects can be similar to each other's over a continuous scale. This explains why subjects can be more consistent or faster in classifying some objects as X compared to others who are less typical X (Rosch & Mervis, 1975). For instance, most of the mammals we see are furry, earthly animals, so dogs are very typical mammals while platypuses aren't. This explains why classifying a dog as a mammal might be easier than classifying a platypus as a mammal (even if platypuses *are* mammals).

There exist two principal understandings of similarity-based classification that are starkly opposed in the literature. "Exemplar" views advance that we compare the to-be-classified stimulus to *all* memorized representations we have for some tested class. For instance, if we need to perceptually classify the picture of an animal as that of a dog or that of a cat, exemplar views argue that we compare the test stimulus to *all* representations of cats we have stored, compute a cat-similarity score, and to *all* representations of dogs we have stored, and compute a dog-similarity score. The highest similarity-score will then drive the classification decision.

"Prototype" views maintain that we assess the similarity-score of the target stimulus to some kind of *averaged* representation. For instance, we compute the similarity-score between the test picture and some stored prototypical representation of a cat, and to some stored prototypical representation of a dog. Again, the highest similarity score drives the classification decision. Prototype representations are "summary" representations. Each of the prototype's dimension value is an average of the dimension values of the class or individual instances the subject has already seen. Encountering new instances of this class or individual thus updates the prototype. Prototype views also predict fuzziness and typicality effects in classification: similarity-scores of a representation relative to a prototype can continuously vary, and, by definition, the more "typical" a representation will be, the more similar it will be to a prototypical representation.

What are the prospects for similarity-based classification processes to count as relevantly *perceptual*? I think they are thin. The main issue is the following: both exemplar and prototype views most often assume that similarity is assessed between an occurrent perceptual representation and prototypical feature descriptions corresponding to some concept stored in semantic memory. The rationale behind such idea is that similarity-based classification without attribution of a specific concept would be semantically meaningless. Without it, one would just be representing that some representation is similar to some other representation(s).

Semantically meaningful classification requires something more. It requires that one takes a decision as to whether similarity to prototype or exemplars is enough to yield the application of a concept to the occurrent representation. The goal of similarity-based views of classification is not merely to argue that we represent some perceived particular as similar to stored representations. It is to explain how such similarity assessments allow us to apply concepts to the occurrent percept.

There are thus three steps to perceptual classification according to similarity-based views: (1) one extracts a perceptual representation of an object; (2) one (sub-personally) assesses the similarity of this representation to representations stored in memory; (3) following a specific decision criterion (e.g., overcoming a minimal similarity threshold), one attributes the conceptual label of the memory representation (as stored in semantic hubs) to which the occurrent representation is most similar to.

Step (3) puts similarity-based classification strategies out of the purview of being a genuinely *perceptual* mechanism since it violates the criterion of non-conceptuality and independence from semantic memory stores. Similarity-based are unlike rule-base views in that subjects test whether a perceived object corresponds to a concept by retrieving only *perceptual*, *modally-specific* information about that tested concept. In other words, similarity-based views see classification as depending on checking the similarity of the current representation with observational retrievals of information. But the goal of this similarity-*based* classification is not to only check for similarity. It is to apply a general concept (that points to observational information, but also verbal information). When I come to classify an object as a cat based on its similarity to some perceptual memories, I attribute the concept to which these perceptual memories are bound. Once this concept is applied, one can also potentially access non-observational information. One has *categorized* the object as a CAT.

Furthermore, many similarity-based views obviously contravene to them being understandable as relying on *perceptual* mechanism by allowing in step (1) and (2) attributes that are clearly not perceptual. For instance, in the celebrated experiments of Eleanor Rosch and collaborators on prototype models of classification (Rosch & Mervis, 1975), it is clear that what they consider as the attributes underwriting family resemblance oscillates between low-level features that are perceivable ("has legs") and abstract features that are not perceivable ("you drive it", "you eat it")[43]. Likewise, in another foundational article where psychologist Amos Tversky proposes a formalized theory for how to calculate similarity to exemplars, he accepts both perceptual features (e.g., the facial traits of faces) and non-perceptual features (e.g., "being a Communist country") as possible features to be considered in a similarity assessment.

Thus, I think similarity-based views of classification, just like rule-based views, describe a mechanism of classification that relies on access to concepts stored in semantic

[43] At least, they couldn't be considered as perceivable without begging the admissible content debate in favor of perceptual Liberalism.

hubs. Similarity-based views diverge from rule-based views only insofar as they describe different processes for retrieving concepts to be applied[44]. The most popular theories of "perceptual" classification are thus really theories of hybrid perceptual-cognitive *categorization*.

III.2.3. Neuroscientific and Computational Models

I think the neuroscientific and computer modelling literatures on object "recognition" substantiates the same conclusion as the preceding section. Terminology should again be manipulated with care. My understanding of recognition as a purely perceptual and thus nonconceptual type of classification is to be contrasted with the very general understanding of "recognition" in the neuroscientific and computer vision literatures as any kind of classification of visual inputs.

Most contemporary neural and computational models of perceptual classification are in line with similarity-based views of classification as described by cognitive sciences. There is a widespread suspicion that it is just hopeless to find a fixed knowledge database to implement object classification (except for very restricted domains of application that are well defined). This suspicion is explicit in the historical abandonment of "classical AI" systems – systems which operated over fixed rules and knowledge databases implemented in a top-down way by scientists) – in favor of dynamic "deep neural network" systems –

---

[44] This might be slightly exaggerated. Actually, there are psychologists who accept some kind of "purely perceptual" and "nonconceptual" similarity-based categorization processes. This is especially true of some developmental psychologists who try to make sense of very young infants' capacities to categorize objects without concepts. Developmental psychologist Jean M. Mandler is a good illustration of someone drawing such distinction between "perceptual categories" and "conceptual categories". She writes:

> [P]erceptual categories work on different kinds of information than conceptual categories. This first difference may be considered a matter of definition. There is an important distinction to be made between people's summary representations of what things look like and their summary representations of what things are. Perceptual categorization computes perceptual similarity. At least early in infancy it does so independently of knowledge about function or kind; indeed, it can occur even in the complete absence of meaningfulness. We could even say that it is not categorization at all but perceptual schema formation, reserving the term categorization for conceptual categories. (Mandler, 2000, p. 30)

This quote is representative of a general idea within the psychological literature, which is that as long as classification is not accompanied by concept application, then it remains a kind of meaningless identification of familiar stimuli, and we can thus put to doubt the idea that it's a genuine kind of classification. Many developmental psychologists actually even deny that there exists anything like a purely perceptual categorization as distinguished from conceptual application. They instead advocate for a continuum between categorizations in early infancy that attributes concepts only through perceptual information, and categorizations that attribute concepts through both perceptual and non-perceptual information (Madole & Oakes, 1999; Quinn & Eimas, 1997, 2000). For these authors, it's thus simply that infants come to attach non-observational (verbal) information to the concepts they are able to apply. All in all then, the idea of a purely perceptual, nonconceptual form of meaningful classification receives bad press in the psychological community. I try to show in the next chapter that actually there *is* some classification that is purely nonconceptual, rehabilitating the meaningfulness of Mandler-like "perceptual categories" (which I call "aspects").

systems which themselves learn from vast amount of data which information and procedures are relevant to complete a given task, such as object classification.

Deep neural networks models for object recognition parallel the same general steps as similarity-based psychological models of perceptual categorization (Battleday et al., 2020; Peters & Kriegeskorte, 2021; Sorscher et al., 2022). In its first layer, the system extracts some features of the stimulus. Then, in deeper layers, some feature information is diminished or simply pruned out while some other is amplified through a differentiation of connection weights. Finally, in the final readout layer, the weighted information is connected to nods which are specifically labelled. The nod having the greatest activity wins out the competition and its label is attributed to the visual input. Through a reinforcement learning procedure (the network trying to minimize error), the connection weights of the network adjust until global network activity maximally activates the correct readout nod for the presented image.

This process can be understood as a kind of similarity assessment: a trained network should activate in the same way (have the same pattern of weighted nod activity) for inputs activating the same readout nod (for inputs yielding the same classification). In a connectionist spirit, we can view the structure of weights of the network as being an implicit representation of previously seen images. For instance, after successful training, when a deep neural network is presented with a picture of a cat, the network activates a weighted pattern that reliably activates the "cat" readout nod. The specific weighted activity pattern for the cat image implicitly indicates that the image is similar to previously processed cat images that shaped the weight structure of the network in a specific way.

Such deep neural network models are increasingly recognized as being good models for understanding the neural implementation of object classification (Spoerer et al., 2017). Neurons are easily thought of as nods in a hierarchical network. Nod activation levels are translatable as some physiological variable (usually: firing rate) and connection weights adjustments can be viewed as a case of synaptic plasticity. Deep neural networks only make use of perceptual information to train their weight structure. They are only sensitive to the superficial properties of images, and they have no knowledge base regarding things like the origin or the function of objects.

Nonetheless, I think these networks ultimately perform a kind of *concept application*: their readout layer is *labeled*, and it is basically telling us whether an image corresponds to some predefined concept, as determined by a labeled training set. At a minimum, the network implements conceptual application in the sense that human coders have conceptualized (labeled) a training set so as to bootstrap the learning procedure. The machine then learns to apply the same conceptual labels as were given by the human coders in the training set. More elaborated deep neural networks go beyond this mere "labeling" approach and also implement language-like semantic networks on top of visual classification networks. These networks exhibit better object classification performances and their activity patterns are more akin to what is observed in fMRI data (Devereux et al.,

2018). These neural networks clearly perform *categorization* (concept-application). They are not good models for a purely perceptual-nonconceptual kind of classification.

At this point, one might be tempted to argue though that neuroscientific results with human and primate brains demonstrate that categorical information can be decoded from purely *visual* brain areas like the visual temporal cortex (Proklova et al., 2016; Schindler & Bartels, 2016; Zeman et al., 2020), thus demonstrating that object classification can be purely perceptual. The problem though is that it is not obvious that such finding really indicates genuine classification or really track purely perceptual processing.

We can find in perceptual neuroscience a debate that mirrors the debate in developmental psychology between "perceptual categories" and "conceptual categories" (Bracci et al., 2017; Bracci & Op De Beeck, 2023). In one camp, we find "neural Conservatives", who maintain that high-level perceptual areas only select some relevant low-level (or "mid-level") contents that are only diagnostic of some high-level content, but maintain that these regions do not in themselves categorize stimuli into high-level conceptual categories (Coggan et al., 2016; Long et al., 2017; Watson et al., 2016; Zachariou et al., 2018). The representation of genuinely high-level conceptual contents is left to further, non-perceptual brain processes.

In the opposing camp, we find "neural Liberals", who defend that high-level perceptual areas are not only sensitive to specific low-level properties, but also classify stimuli into high-level conceptual categories (A. C. Connolly et al., 2012; DiCarlo et al., 2012; Huth et al., 2012; Riesenhuber & Poggio, 2002; Schindler & Bartels, 2016). But we should note straightaway that if we interpret ventral temporal cortex activity as performing some kind of conceptual application, this will jeopardize our understanding of this brain area as being purely perceptual. I think the conclusion we should draw if it turns out that the ventral temporal cortex (or other higher-level "perceptual" areas) accesses some semantic memory stores and perform some kind of concept application itself is that such area is a composite, hybrid perceptual-cognitive area. This area would both be extracting and selecting low-level features, and it would access a database of "semanticized" labels (i.e., concepts) to apply to the visual input. Actually, as is often found in the literature (Romanski & Chafee, 2021), it might be more appropriate to talk of the ventral temporal cortex as an "association" or an "integrative" visual-conceptual area (Martin et al., 2018), rather than of high-level perception *tout court*.

It thus seems that the interpretation of the neuroscientific or computer modelling literatures regarding perceptual classification leads to a dilemma akin to that already encountered with psychological models. On the one hand, one might be tempted to understand object classification as some kind of concept application mechanism (empirical and modelling results strongly supporting such view). But in that case, the horn to face is that we are not dealing anymore with a purely *perceptual* classification mechanism, but we are describing something like a hybrid perceptual-cognitive categorization mechanism. Talking of high-level "vision" or deep neural networks as performing *perceptual* classification proper is thus representationally misleading.

On the other hand, one might instead choose to view object classification in high-level perceptual areas or deep neural networks as not connected to any kind of concept application but rather as performing some purely perceptual classification through mere selective feature sensitivity. But in that case, the horn to face is that it is not certain that we are dealing with anything like meaningful high-level perceptual *classification* (representing something as in some equivalence class) anymore (cf. fn. 44). Talking of the ventral temporal stream or deep neural networks as performing meaningful perceptual *classification* on this view might thus be misleading.

Recognition of this dilemma shows how difficult the quest for a plausible *perceptual* mechanism for high-level *classification* is. Keeping this issue implicit by not advancing a plausible psychological mechanism to ground high-level contents in perception can only make the debate more insoluble. But clairvoyance about one's disease is a first step towards healing. Now that such clarification of the dilemma is done, we can begin to find a solution. Such solution lies in facing the second horn of the dilemma: a Liberal should be able to maintain a rejection of the necessity of concept application processes for perceptual classification, and keep alive the possibility of meaningful *nonconceptual* classification. This is exactly what the next chapter aims at showing.

IV. SCHEMATIZATION

The previous chapter showed us that, as the scientific literature stands, the proposed mechanisms for *perceptual classification* (as studied in cognitive psychology, neurosciences, or computer vision) rely at some point on some access to semantic stores. Such access contravenes the idea that such classification is purely *nonconceptual*, and thus is not able to justify *perceptual* Liberalism. Nonetheless, as I now defend, *similarity-based* classification gave us enough empirical material to defend a primitive kind of high-level content in perception, one that doesn't depend on concept application and access to semantic memory, but that uses only perceptual, i.e., nonconceptual, processes.

The main idea behind similarity-based classification is that one compares an occurrent percept to representations stored in long-term memory (either prototypes or exemplars), and then applies to the percept the concept linked to the representation that is most similar to it. Such similarity-based conceptual application might uncontroversially be seen as yielding a kind of high-level representation, though it is one that cannot be viewed as exclusively perceptual but comes down to some kind of post-perceptual judgment.

Still, there is a processing stage before concept application that is of interest to us: the similarity assessment itself. We can see things as being similar to things we have seen before. We do so effortlessly and automatically. If I cross someone in the street who looks very much like Donal Trump, I cannot help but see this person as <u>being similar to Donald Trump</u>. This is so even if I'm not voluntarily trying to retrieve from semantic stores *who* this person resembles (i.e., to <u>DONALD TRUMP</u>). And I persist in seeing someone as <u>looking like Donald Trump</u> even if I know very well that '<u>this is not DONALD TRUMP</u>'.

This dissociation stands as a good prima facie reason to think that something nonconceptual and purely perceptual is going on in such a recognition event. Of course, there can also be something *conceptual* that is further going on: I could see that '<u>this person looks like DONALD TRUMP</u>', or simply see that '<u>this person is DONALD TRUMP</u>' (if I have no defeating reasons to think that this person is not Donald Trump). This latter representation would be a case of a similarity-based *categorization* (a demonstrative post-perceptual thought).

Notice though the fundamental point that the similarity assessment itself is not identical to the conceptual application (categorization). The similarity assessment happens *prior* to the categorization (after all, the latter is only similarity-*based*). A good illustration of this idea comes from the fact that one might be sensitive to a certain similarity (e.g., see someone as <u>looking like Donald Trump</u>), but fail to retrieve from semantic memory the

particular individual that corresponds to this similarity (see *that* 'this person is DONALD TRUMP').

Now, we can wonder whether the similarity assessment itself (*without conceptual application*) might be sufficient to constitute a kind of high-level *nonconceptual* classification in perception. Two obstacles stand against such an idea. First, we need an argument as to how similarity representation on its own can yield a kind of high-level representation. One could argue that, without conceptual application, one could just perceptually represent some similarity with respect to some low-level features, which might itself be taken as some kind of low-level representation. This low-level representation represents only *meaningless resemblance*, a property that is not sufficient to ground perceptual Liberalism. After all, even a Conservative could reasonably accept that we perceive similarity between things we perceive. What we need is similarity grounding some kind of *meaningful equivalence* between representations to ground a genuine kind of high-level classification.

Second, we need to be careful that the mechanism that gets us a similarity representation does not bring through the back door non-perceptual, conceptual elements. We need a *purely perceptual* way of getting a similarity assessment. A Conservative might here see an opportunity to resist: is similarity really providable through purely perceptual means; or is similarity dependent on concept application? Can one really see two particulars as similar without thinking about what makes them similar? Can one really see something as looking like a cat without post-perceptually judging that 'it is a CAT'?

These obstacles can be overcome. I'll show that the mechanism through which one gets to a similarity assessment, *schematization*, is (1) purely perceptual, and (2) grounds a primitive kind of high-level representation: *aspects*. This chapter is dedicated to point (1). I will argue that, to produce similarity assessments, a psychological mechanism must answer "Goodman's challenge", or the issue that similarity is an indetermined relation if it is not relativized to some *respects* (IV.1). I show that we can identify a mechanism, "structuring", which precisely solves such respect-indetermination problem in perception. Once structuring is in place, I argue perception can produce a similarity priming process. I call such conjunction of structuring and similarity priming processes "schematization" (IV.2). I next present theoretical and empirical arguments for the existence of schematization processes in perception (IV.3). Schematization is the purely recognitional, non-categorical, non-conceptual process of classification that we are looking for to vindicate perceptual Liberalism.

IV.1. GOODMAN'S CHALLENGE

Imagine one wants to compare the similarity of two cars C_1 and C_2 to determine whether C_2 is the same model as C_1, which is known to be a Fiat 500. Clearly, there are some features of our perceptual representations of C_1 and C_2 that are completely irrelevant as to whether they are similar regarding the model of the car. The *position* of the car is for example

irrelevant as a respect for model similarity: seeing C_1 on the left and C_2 on the right is not a difference that should matter. The same goes for many perceptible features: it's irrelevant whether one car is red and the other green; whether one car is brightly illuminated and the other dimly illuminated. This point is even more evident if we go beyond features that are expressible in natural language and consider representational features as described in the neuroscience of vision (e.g., fine-grained local contrasts): many, if not most, of these features will be irrelevant for similarity assessments.

The number of irrelevant features is *even more* expanded if one lets in non-perceivable features: the fact that both cars are owned by Caroline is irrelevant to their similarity with respect to their model. To assess their similarity with respect to car-model, one must then select some appropriate feature dimensions to make the similarity assessment. In our example, probably one of the most relevant *perceivable* features will be the shape of the cars, and a relevant non-perceivable feature will be whether these cars were assembled in the same factory. Notice that obviously, perceptual similarity is a defeasible criterion for model classification: one might judge that C_2 is a Fiat 500 because it has the same shape as C_1, but then one might also learn that C_2 is just a counterfeit that was assembled outside of a Fiat factory. Similarity-based classification is amenable to corrections by other kinds of information.

At this point, I must make a crucial digression to rebut an enduring philosophical form of skepticism regarding such relativized similarity relations, a skepticism famously upheld by philosopher Nelson Goodman (1972) and against which much of the similarity-based classification psychological literature can be read. Goodman thought that the respect-dependency of similarity assessments made similarity explanatorily irrelevant to explain many mental phenomena, and he explicitly claimed that "dyadic likeness between particulars will not serve to define those classes of particulars that have a common quality throughout" (Goodman, 1972, p. 443). In other words, he denied that similarity could ground any kind of classification. This skepticism is grounded in the idea that, if similarity depends on circumstantial respects that determine the properties to be compared, then similarity will just be redundant information. For Goodman, "to say that two things are similar in having a specified property in common is to say nothing more than that they have that property in common" (Goodman, 1972, p. 445).

I think this skepticism is largely overblown. Goodman assumes that similarity is a matter of sharing features. Thus, two objects do or do not share some feature *x*. But this view of similarity is unwarranted (at least in the context of psychological theory). Similarity can instead be understood as the inverse of the *representational distance* between two representations. Take **Figure 3** below. One can say that A and B are similar with respect to color even if strictly speaking they do not display the same shade of red. Likewise, B and D are similar with respect to shape even though they are not the exact same ellipses. The existence of similarity relations between these objects is a non-trivial psychological fact. It reflects an important property of perception: that it is structured, and that it represents

feature values in some ordered space[45] yielding determined representational similarity relations (Churchland, 2012; Clark, 1992; Gärdenfors, 2000; Matthen, 2007 all explore the representational structuration of perception in various ways). Colors for instance are thought to be ordered in a *color space*, a space in which orange and red are color values that are representationally closer (are more similar) than red and blue. Notice the difference here with *words*: the words "red" and "orange" are no more similar that the words "red" and "blue"[46].

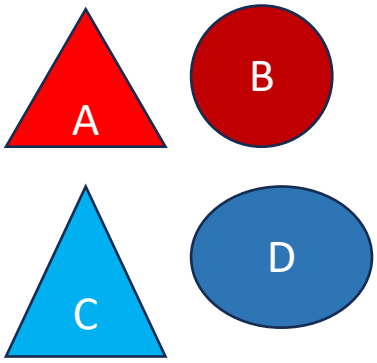

Figure 3. Two possible similarity assessments.

Pace Goodman, perceiving the similarity between two objects is not akin to judging the number of their common features. Being geometrically organized, perceptual dimensions give us a distance metric between their values. We can call such distance on single, separable, low-level perceptual dimensions a *primitive similarity relation*. Color shades for instance bear a primitive kind of similarity that we can perceive. The same holds for other low-level perceptual dimensions: brightness, shapes, motions etc. I wish to insist here that

[45] This space is of course only representational. Similarity (closeness) in the representational space does not (necessarily) amount to anything like a physical similarity (though representational might be coded as similarity in brain activation patterns).

[46] The idea that perception is geometrically structured and sustains relations of similarity between representations has led many philosophers to argue that the representational *format* of perception must be different from that of language and language-like thought. In particular, it has led many to argue that perceptual representations must bear a *continuous analogue* relations (or a kind of "isomorphism") with the structure of the environment (Beck, 2019; Block, 2023; Shepard & Chipman, 1970), which might indicate that perceptual representations have an *iconic*, and not *discursive*, format (Fodor, 2007; Lyons, 2022; Quilty-Dunn, 2017). In this thesis, I prefer to refrain from entering into the *format* debate about perception, preferring to focus on the *content* (conceptual v. nonconceptual) debate. I tend to agree with Block that nonconceptual content and iconic formats probably bear some important relations, though I don't think my defense of nonconceptual contents in perception is necessarily bound to accepting an *iconic* format for them.

one should keep in mind that such primitive similarity relation only expresses the fact that perceptual representations bear a certain distance in perceptual representational space. It does *not* express the idea that we *explicitly represent such distance*. We do *not* have a meta-representational capacity to represent *similarity* as such. We simply represent perceptual values in a structured way that constitutes perceptible distance relations.

Now, Goodman could argue that such primitive similarity relations are insufficient to determine whether the "dyadic likeness" of two objects is sufficient or not to classify them together based on similarity (subsuming them under the same sortal, like "red"). In fact, dyadic similarity (similarity between two particulars) is too unstable, being dependent on circumstantial and volatile respects. But again, this worry is unwarranted. What one needs to classify particulars together based on similarity is not some kind of absolute "dyadic likeness" value between two particulars. What one needs is merely to determine whether two particulars are more like each other *than to a third one* (Gauker, 2011, Chapter 7). For instance, taking **Figure 3** again, if one knows that B is an X, and C is a Y, then one needs only to determine whether A is more similar to B than to C or *vice versa* to determine whether A is an X or Y (imagining that A is necessarily either an X or a Y).

Of course, such similarity assessment will change depending on the respects to which we compare objects (A is more similar to B than to C with respect to color, but more similar to C than to B with respect to shape). But this is not a problem since we are not interested in *absolute* similarity, but in the *relative* similarity of A to determine to which other object it is most similar to *with respect to some specific dimension.* That similarity is relative to some respects doesn't impede its psychological import because *primitive* similarity relation is a non-reducible psychological fact. That red is more similar to orange than to blue is a primitive psychological fact under the color dimension.

Goodman's fundamental intuition should thus be taken seriously: there is no similarity without selecting some specific respects for similarity assessments. Similarity-based theories of classification thus must suppose that there exist mental processes that select those respects for similarity. But this idea might generate a difficulty: is it possible to determine respects for similarity through purely perceptual processes? Can one determine respects for similarity in a purely *nonconceptual* way? Can one see two cats as similar without possessing a CAT concept? I will name this issue "Goodman's challenge". Of course, the challenge is powered by the view that classification (e.g., of an object as being a CAT) is thought to be dependent on categorical, concept-application processes (Deroy, 2019). The challenge for us is to find forms of respect selection that are not dependent on concept applications. I now propose that the process of *perceptual structuring* can fulfil such a challenge.

IV.2. Solving Goodman's Challenge through Schematization

IV.2.1. Perceptual Structuring: Seeing in Respects

We can get the intuition of perception being able to nonconceptually select respects for similarity by considering **Figure 3** again. When looking at it (and imagining we fix our gaze to the central cross), we can attend differentially to the features of the figure. If one attends to color, one will see the pairs A-B and C-D as being perceptually similar to each other. If one attends to shape, one will see the pairs A-C and B-D as being similar to each other. Phenomenologically speaking, this seems like a purely perceptual phenomenon: it's not that we *think* that A-B and C-D form pairs of similar objects because they have resembling colors. Rather, we *see* this directly through some kind of Gestaltic effect: A-B and C-D are visually grouped together when we attend to color (and A-C and B-D when attending to shape). Thus, feature-based attentional selection offers a good illustration of a *perceptual* mechanism playing the role of respect-selector for similarity assessments.

I take feature-based attentional selection to be just an instance of a much more general phenomenon of dimensional selection that I call *perceptual structuring*. I take "structuring" very generally to be a perceptual process by which one informationally *simplifies* and *organizes* a previously more informationally complex and "disorganized" representation. Structuring can take the form of making a dimension of a perceptual representation more *salient* than others (as in attention), or fusing several represented features into a more general one, or simply discarding some perceptual dimensions. All these processes result in some kind of information simplification or organization of a perceptual representation. I leave the definition intentionally as rather vague and metaphorical. I don't think structuring is anything like a natural mental kind. It rather subsumes many different processes as described in vision sciences. What is important is that these variegated structuring processes conspire to produce a perceptual representation that does away with some feature information and emphasizes (or makes salient) other information.

"Emphasis" or "saliency" is to be understood as being representationally prioritized in subsequent psychological processing[47], among which, crucially, we can put similarity assessments. Notice that through structuring, it is expected that making some perceptual dimension more salient should naturally make less salient some other dimensions. Structuring should thus be understood against the background of a (theoretical[48]) "unstructured" representation, i.e., a perceptual representation that represents a maximum number of perceptual features and that doesn't emphasize or prioritize any dimension over some other.

---

[47] I thus do *not* take structuring to be constitutively a phenomenological phenomenon, though the representational result of structuring *can* yield a specific phenomenology (as in cases of structuring through *attending*).

[48] I say that this unstructured representation is merely theoretical because it's possible that some structuring processes are always-already at work in any perceptual representations.

The idea of some sub-personal, bottom-up, automatic structuring of some unstructured perceptual representation is at least as old as the emergence of computational understandings of vision (Marr, 1982). It is at the heart of all up-to-date deep neural network models of vision, which core algorithmic goal is to transform and simplify an overly complex input representation so as to keep only with relevant features for various tasks (among which *classification* is the most widely investigated). One of the most important and best described form of perceptual structuring is the organization of legions of local edge representations into elementary contours (Marr, 1982), object fragments (Ullman et al., 2016), or global shape (Ayzenberg & Behrmann, 2022).

Attention can be understood as some personal-level, top-down, voluntary modulation of bottom-up perceptual structuring. There are many cases of perceptual structuring that are *not* dependent on attention, and that are *not* effects of personal-level, top-down, voluntary modulations of perception. Much perceptual structuring is a sub-personal, bottom-up, automatic mechanism as described in vision sciences and artificial neural network models.

I think understanding perceptual structuring is crucial to get to understand how anything like *nonconceptual similarity assessment*, and thus ultimately *high-level perceptual contents*, are possible. I thus dwell a bit on the description of such perceptual mechanisms.

The idea of structuring lies at the core of many models of object classification in neuroscience, where it is diversly known as "dimensional reduction", "sparsification", "untangling", or "filtering". Untangling has been upheld as a solution to the "problem of invariance" in the object classification literature (Biederman, 1987; DiCarlo et al., 2012; DiCarlo & Cox, 2007; Riesenhuber & Poggio, 2002). What is meant by the problem of invariance is actually treading on ideas we encountered describing Goodman's challenge: for perceptual representations to serve as useful inputs in classification process, they must be insensitive (invariant) to features that are irrelevant to the classification process, so as to keep only with relevant features.

As we saw, position is most often an irrelevant feature dimension to classify an object. The position of an object we could say is *garbage data* for classification ("garbage in, garbage out" as computer scientists say). To illustrate this idea, let's imagine a perceptual space, where each axis corresponds to a feature dimension (**Figure 4**). Each representation of a particular can be visualized as occupying a specific point in this space, depending on the values it has for the corresponding dimensions. Distance between each points in this space can be viewed as representing the inverse of the similarity between two representations (i.e., dissimilarity).

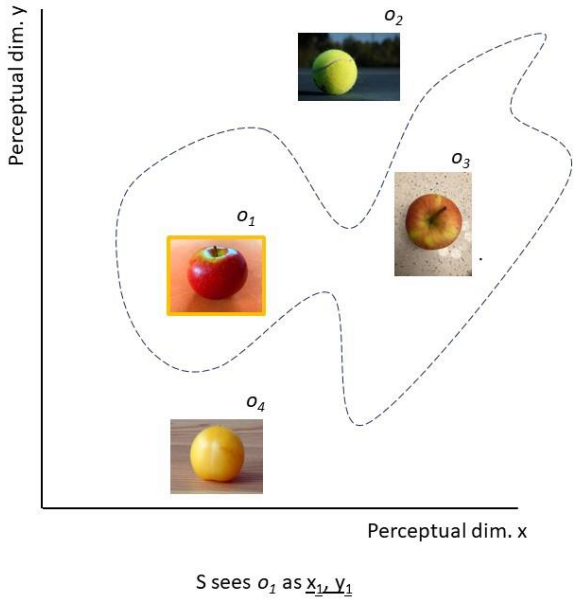

Figure 4. An unstructured, semantically tangled perceptual space.

The problem of invariance can be easily visualized from such space: the problem is that similarity relations in an *unstructured* representational space are *semantically meaningless*. If one takes the closest representations to an unstructured representation of a particular, one will most probably get a representation of a particular that is semantically unrelated to the first, *because these particulars will be similar with respect to irrelevant ("garbage") feature dimensions*.

Take **Figure 4**, which represents an unstructured perceptual space for four particulars: two apples (o_1 and o_3), one tennis ball (o_2) and one plum fruit (o_4). If we consider o_1, we can see that it is most similar to o_3, another apple, with respect to the y-dimension (one can imagine that this dimension represents *shape* for instance). But we can also see that o_1 is more similar to o_4 with respect to the x-dimension (which could be for instance the irrelevant dimension of *position*). In this unstructured representational space, similarity relations aren't meaningful, i.e., they aren't similarity relations that could correspond to some relevant class (*apple*, *plum*, *tennis ball*). In such space, we are really at the mercy of Goodmanian-style argument stating that similarity as such is semantically meaningless and psychologically irrelevant for classification.

Another way to visualize the idea that meaningful similarity assessments are not possible in such a space is by thinking in terms of "manifolds": manifolds represent the possible space potentially occupied by either one and the same particular, or particulars from the same class. In **Figure 4**, I have represented the manifold for all *apples* in the

unstructured perceptual space by the space delimited by the dotted line. o_1 and o_3 fall within this manifold. What is important to notice is that, in such space, the manifold for all apples is *non-convex*: there are many representations that are potentially in-between two apples that are not apples (see Gärdenfors, 2000 for some mathematical formalizations of the notion of *convexity* in representational spaces). In other words, something could be more similar to a particular apple than another apple, *which is not an apple*.

If we wanted to draw a boundary between apples and non-apples in such unstructured space, this boundary would need to be non-linear. One would need to know a very complex rule (the one that draws the non-linear boundary between things that look like apples and things that don't) to make a similarity assessment with respect to apple-ness. James DiCarlo and collaborators, who introduced the notion of "manifolds" in the object classification literature (DiCarlo & Cox, 2007), have pointed out that it is computationally implausible that any object classification algorithm (including the one running in the human mind) could learn such complex classification rule, i.e., could know how to distinguish such complex, non-convex manifolds. DiCarlo takes the example of the manifold of a particular person's face ("Sam"), generated artificially by a computer program. Sam's face can occupy many points in the program's image space. To simplify, the program limits differences in Sam's face images to its orientation (cf. **Figure 5** below, top-left). Of course, in real life, the manifold corresponding to Sam's face would be much more extended and complex, because it would also include differences in illumination, size, facial expressions etc.

Now, if we want to code a program that is able to distinguish images of Sam's face from images of Joe's face, DiCarlo et al. defend that that would be equivalent to the program being able to compute a plane that strictly separates between the manifolds of Joe's and Sam's face (**Figure 5**, top-right). But problematically, when one codes these manifolds in terms simply of the pixel values of the images, we see that the manifolds are heavily *crumpled* or *tangled* (**Figure 5**, bottom-right). Any possible hyperplane would cut across the manifolds corresponding to both Joe and Sam, making their recognition very difficult, if not impossible. In such an unstructured pixel space, one thus *cannot* computationally distinguish between Joe's and Sam's face. One thus needs to build a "good" space where precisely there exists a strict orthogonal separation between targeted manifolds (as in **Figure 5**, top-right). As DiCarlo et al. say, one needs to *untangle* the manifolds by building a good space in which these manifolds are separable by a (hyper)plane. This is done by finding relevant dimensions of invariance, i.e., dimensions of variation to which the neural space should stop responding (for instance, *position*). Brute reduction of dimensions by itself is insufficient though since many untangled spaces still do not yield appropriate separations between manifolds (**Figure 5**, bottom-left). One must *learn how to untangle* (and perceptually structure in general), a problem I will come back to in VI. KNOWING HOW TO Schematize.

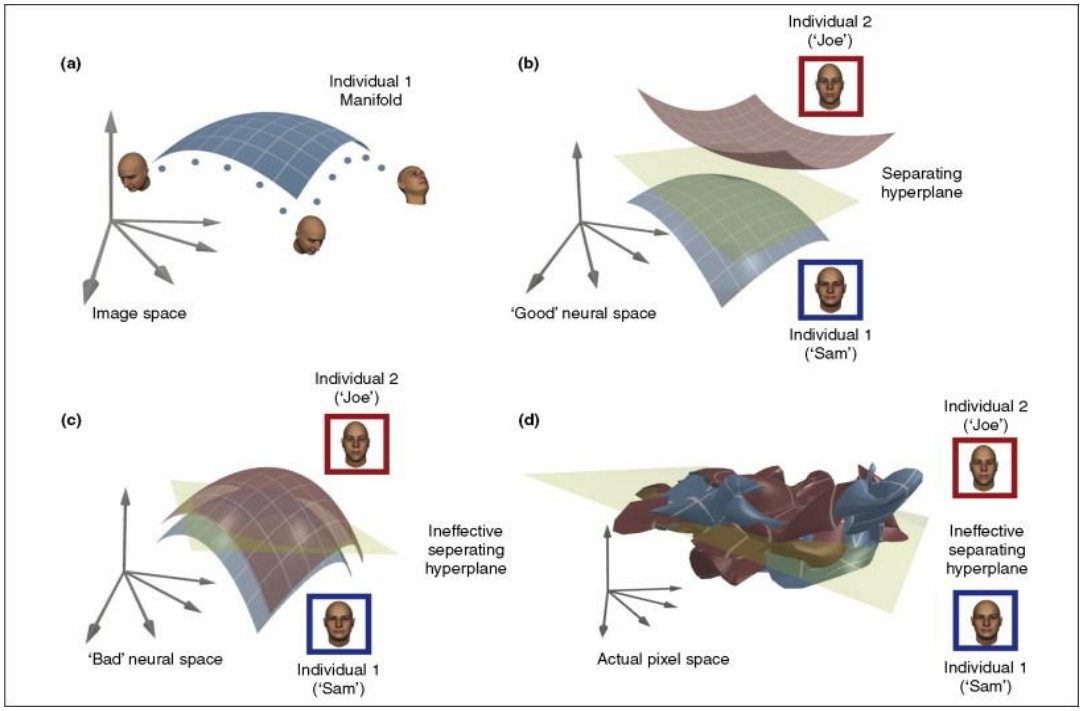

Figure 5. Illustration of manifold untangling.

From (DiCarlo & Cox, 2007). Permission to reuse from the publisher. Copyright © 2007 Elsevier Ltd.

Importantly, DiCarlo et al. (DiCarlo et al., 2012; DiCarlo & Cox, 2007), and others after them (Sorscher et al., 2022), have demonstrated that the computational process of manifold untangling through projection of one (tangled) representational space into another, simplified (untangled) space is the best computational model of the neural activity of the ventral temporal pathway. As DiCarlo et al. conclude, "[in] sum, the experimental evidence suggests that the ventral stream transformation (culminating in IT) solves object recognition by untangling object manifolds" (DiCarlo & Cox, 2007, p. 337).

Untangling remains only a useful *model* though, and we can legitimately wonder whether full-fledged *invariance*, and thus untangling understood as feature *discarding* really occurs in biological systems like human vision. It is empirically established that higher-level perceptual areas in the ventral and dorsal streams remain sensitive to many low-level features (Ayzenberg & Behrmann, 2022; Kravitz et al., 2013). Biologically speaking, it might thus be more correct to view the visual system as *deemphasizing* certain dimensions (and emphasizing others), instead of literally taking it as *eliminating* irrelevant dimensions of variation. In biologically realistic untangling, the perceptual system sub-personally gives more computational priority to some perceptual dimensions than others *so as to produce meaningful manifolds in the untangled perceptual space*.

Much of such structuring processes as described in the creation of "untangled" spaces relies on the *fusing* or *binding* of low-level spatial, local features of objects. So, for instance, local edge shape representations (e.g., representation of the orientation and curvature of part of the ear) are irrelevant for creating distinguished manifold representations corresponding to Joe's face and Sam's face. Keeping with such local feature information tangles the manifolds corresponding to these objects. To untangle these manifolds, one needs to become insensitive to local edge shape representations. One way this can be realized is by bounding local edge shape representations into a unique representation of a global shape. By doing so, one becomes insensitive to local edge shape information, and is sensitized to a new feature dimension: global shape. On thus performs a kind of many-to-one dimensional reduction. Such new dimension might sufficiently untangle the manifolds corresponding to Joe's face and Sam's face, as in **Figure 5**, top-right).

Again, this idea of some untangling process through fusing or "pooling" local features into more global ones is at the heart of hierarchical models of vision since Marr's time (Marr, 1982), and they have gained much traction in up-to-date deep neural network models of object classification (Spoerer et al., 2017). Transformations of local spatial representations into more global spatial representations of contours, parts, and shapes is by far the most important structuring principle for object classification.

Most of perceptual structuring happens automatically through bottom-up, feedforward computational processing performing some kind of untangling (through dimensionality reduction and processing selectivity), a phenomenon classically observed in neuroscience descriptions of the ventral pathway and computer vision models of object classification. Voluntary personal-level attentional selection participates in perceptual structuring precisely insofar as it modulates the untangling process (an idea that I will come back to in VI.2 CONTROL).

What matters for now is to keep in mind the existence of such structuring processes in general, since as I now show they hold the key to solving Goodman's challenge. In fact, structuring provides us with a computationally and neurally plausible solution to face Goodman's challenge. If a perceptual space is properly structured, then the similarity relations for a given representation in this space will be meaningful, since representations in this space should form coherent representational manifolds that are strictly separated. One can thus exploit such coherent similarity relations for classifying objects.

Figure 6 illustrates the idea of perceptual structuring in a simple 2D-to-1D structuring leading to similarity-based APPLE categorization. While the two apple particulars o_1 and o_3 do not cluster together in the unstructured perceptual space (relatively to the ball or the plum), they *do* cluster together if we get rid of dimension x and keep only with dimension y. In this new untangled, structured space, apples and non-apples are convexly separated: there is nothing in-between two apples that is not an apple. They form a coherent manifold. One can easily compute a plane (here: a line) that separates the apple cluster from clusters of non-apples. We can be sure that anything that falls between these

two lines is an apple (or at least, something that *looks like* an apple, as even a plastic apple will fall between these lines).

A structured perceptual space where objects cluster together in a convex way is a space that is easily exploitable by similarity-based categorization processes. In fact, the space to be searched for is simplified and similarity relations in this space should be semantically meaningful, because structuring is not done at random, but is learned so as to subserve human behavior, including classification.

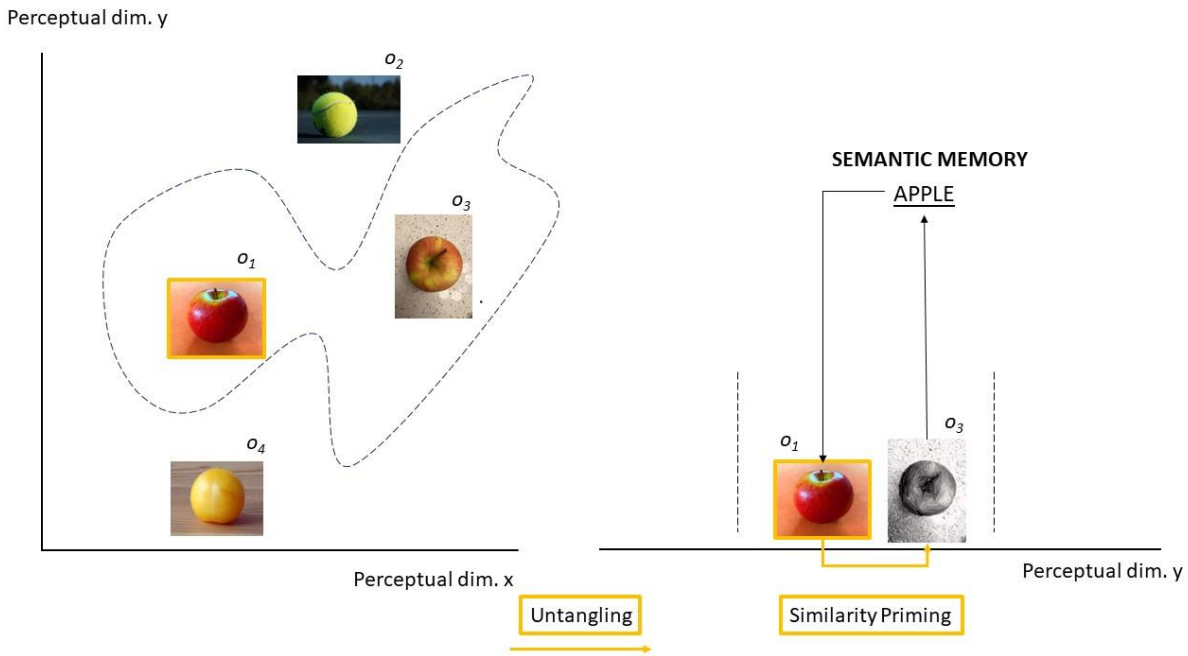

Figure 6. Perceptual structuring leading to similarity-based categorization.

Categorization (concept application) follows naturally: one can retrieve from memory a semanticized (conceptualized) representation that is maximally similar to the representation of o_1 *in the structured space*, and then apply the conceptual label of o_3 (APPLE) to o_1, as illustrated in **Figure 6**. One only needs to retrieve a memory representation that falls in the same manifold. I will describe in more detail the process of similarity retrieval (or "similarity priming") in the following section. For now, I concentrate on the structuring process.

Notice that such meaningful similarity relation is not present in the tangled space: retrieving from memory similar representations to o_1 will yield retrievals that are semantically meaningless. One could irrelevantly retrieve memories corresponding to o_2 (BALL) or o_4 (PLUM). The central idea to retain here is that structuring allows one to create new perceptual spaces in which perceptual representations cluster in a semantically coherent way (in coherent manifolds) *with respect to specific dimensions*. Or to put it differently: structuring allows one to create new perceptual subspaces in which *similarity*

relations are semantically meaningful with respect to specific dimensions. Structuring thus models how similarity-based categorization is even possible and assuages Goodman's challenge.

Of course, the fact that one can represent an object under some specific respects when classifying it does not destroy the low-level feature representations that are taken to be irrelevant. For instance, even if classifying something as an apple based on similarity requires to be representationally insensitive to position or color features, it remains that we still represent the apple's position and color somewhere in the perceptual hierarchy. This is obvious at a phenomenological level. It's also a direct derivation of the fact that one structures by building a *new* representation (in high-level vision or higher-layers of a neural network) that coexists with the low-level representation from which it necessarily derives.

Structuring in high-level perception, as described by cognitive sciences and computer models, is a surprising mechanism when we come to think of it. It is surprising because structuring is a process that doesn't in itself *create* new representational information: it only filters (untangles) and organizes some lower-level information that is overly complex in order to subserve higher-level capacities, such as object classification. We do not have full access to our structured representations. Most structured representations are inaccessible at a conscious level. If we experienced the structured representations that subserve classification, we should often experience positionless, colorless, distanceless objects, which is near to impossible to imagine. Our perceptual experience does not express all that happens in perception. It stands in-between[49] purely unstructured information (which can be thought of as the retinal image) and the most structured information culminating in high-level perception (which preserves only classification-relevant information). We should then beware of not taking perceptual experience as providing us with a transparent image of the entire computational and representational capacities of vision. Again, this is a good reason not to make the admissible contents of perception primarily depend on introspective intuitions about perceptual experience.

Perceptual structuring is thus computationally modeled as a way to simplify a representational space so as to implement some kind of (linear) rule-based classification (DiCarlo et al., 2012; Sorscher et al., 2022). But such perceptual structuring might also just as well serve *similarity-based* classification, by providing a neat solution to Goodman's challenge. In fact, structuring a perceptual representation makes perceptual similarity computable. Structuring *by design* determines respects for similarity by building coherent, distinguished representational manifolds for various classes of objects.

Structuring reformats and streamlines complex low-level perceptual representation to form new simplified representations. Structuring creates a new perceptual

[49] Jesse Prinz's (Prinz, 2012) "Attended Intermediate-level Representations" view of conscious perception proposes potent arguments for this idea that conscious experience is grounded in intermediate-level representations in perceptual systems standing in-between confused early vision and highly structured high-level perception.

representational space in which similarity relations between representations are semantically coherent: close representations in that space are expected to correspond to the same class or individual. Structuring thus provides us with a solution to Goodman's challenge: it makes similarity meaningful *by design* without having to represent explicitly in virtue of which respects two representations are similar.

Nonetheless, *as long as one does not match the structured representation with similar representations*, no similarity-based classification is produced, and thus we cannot talk yet of anything like a *high-level* representation. Structuring is crucial and necessary for such meaningful matching to occur since unstructured representations would yield meaningless similarities. But structuring itself is not performing a similarity computation. A second mechanism is needed: *similarity priming*. Together, structuring and similarity priming form the *schematization* process which I argue *does* form a kind of primitive, nonconceptual, purely perceptual kind of classification.

IV.2.2. Similarity Priming

Manifolds created by perceptual structuring are not represented as such by the subject. Manifolds are theoretical entities that exhaustively represent all the possible positions of an object corresponding to a certain class or individual within a representational space. Obviously, we do not explicitly know all the ways that a cat could look like (i.e., we do not represent the cat perceptual manifold as such). But when we perceive a cat, and represent it through a relevantly structured representation, we can suppose that any representation that is similar to the current one will *also* be one of a cat (cf. **Figure 6** again). A close (similar) representation will pertain to the same manifold. Thus, it would suffice for the perceptual system to retrieve similar representations to the structured current one to find representations that should be semantically coherent with it (providing, again, that structuring process have been learned efficiently). When one sees a cat, one could retrieve similar representations to it under certain respects, and what one should find, if the perceptual structuring is relevant, is some other representations corresponding to a cat. I call such a process of retrieving similar perceptual representations to the current structured percept *similarity priming*.

We can get a sense of the existence of similarity priming intuitively: when we see some particular, we can instantly and automatically come to see it as being similar to other things. I see an animal, and it looks similar to a cat. I see someone, and he looks like Matteo Renzi. The most natural way to think of such similarity priming is in terms of a current representation activating similar *memory traces*. When I see an animal, I produce a structured representation of it, and I can thus automatically activate perceptual representations that are similar to this animal with specific respects (e.g., global shape). These primed representations might be representations of *cats*. The fact that I activate memories of cats when seeing a cat is *not* a consequence of me categorizing the current percept as that of a CAT, and thus retrieving perceptual memories corresponding to my CAT

concept. Rather, it is because I am priming *similar representations in a relevantly structured perceptual space* that what I happen to prime are actually cat memories. If one takes **Figure 6** again, the processing order is the following: first structuring (untangling), *then* similarity priming, *then* categorization. After all, and to repeat, psychologists talk of *similarity-based* categorization, not of *categorically-based* similarity. Similarity priming is computed *before* categorization. It *serves* categorization. The fact that one retrieves representations of cats when seeing a cat is allowed by relevant perceptual structuring, not by categorization.

I prefer to talk of similarity "priming" and not similarity "matching" because I take the retrieval of similar representations to the occurrent structured percept to be *implicit* and *non-declarative*. One shouldn't think of similarity priming as the conscious, explicit recall of similar memories from declarative episodic memory. Similarity retrievals in similarity priming appeal to implicit memory stores, as traditionally described in works on perceptual priming under the term "perceptual representation systems" (PRS) (Schacter, 1990; Tulving & Schacter, 1990). Daniel Schacter describes PRS as follows:

> An important feature of these subsystems, and PRS more generally, is that they process and represent information about the *form* and *structure* of words, objects, and other kinds of stimuli, but do not represent semantic or associative information about them. PRS does, however, have connections with semantic and other systems. (Schacter, 1990, p. 550, *original italics*)

While the term "perceptual representation system" has fallen somewhat out of fashion in the scientific literature, the idea of an implicit, non-declarative, "presemantic" perceptual memory system involved in perceptual priming is still well alive today (Cooke & Bear, 2015; Luck & Hollingworth, 2008), though it is less intensely studied as its declarative and explicit counterparts, episodic and semantic memory. Perceptual priming is said to be "presemantic" in the sense that "(a) it occurs whether or not subjects perform semantic encoding operations, and (b) it is quite sensitive to changes in perceptual properties of target information" (Schacter, 1990, p. 545). Perceptual similarity priming is only driven by perceptual information. This idea of a "presemantic" memory system distinguished from a "semantic" one is crucially in line with my own distinction between "nonconceptual" (or we might say "preconceptual") processes in perception and "conceptual" process in thought *which* point to multiple stores of information (in particular *verbal* systems). Here example, here is Schacter again:

> In the object domain I think you can make somewhat of a sensible distinction if you talk about a semantic domain composed of functional, associative, and perhaps contextual properties of an object. These go beyond the physical form and structure of an object; the presemantic system that I have discussed is restricted to that physical form. (Schacter, 1990, p. 569)

Likewise, similarity priming is "presemantic" (or "nonconceptual") in the vision psychologist's sense that it does not tap into transmodal conceptual information and

remains limited to perceptual, modally-specific information as provided by the current percept and perceptual memory traces.

It is useful to distinguish between three potential kinds of similarity priming: *exemplar similarity priming*, *prototype similarity priming*, and *simulated similarity priming*.

Exemplar similarity priming is akin to the Empiricists' idea of association between perception and perceptual memories. The kind of representations one primes are mnemonic traces: one sees a cat as being similar to some cats one has seen before. Remember that I am taking perceptual memory to not situate representations at a particular time in the past. When I say that one primes similar representations of cats one has seen before, I am not saying that one recalls experiences of cats in the past. One rather primes time-neutral "object memories" (Openshaw, 2022) corresponding to cats one has seen before.

Prototype similarity priming takes on the idea that perception's current representation primes mnemonic traces, except that the kind of memory traces that are considered are not literal traces of past perceptions, but rather they are *averages* of them: one sees a cat as being similar to a *prototypical* cat. Of course, the prototypical cat will be in the cat manifold (it is its barycenter), though the prototype of a class does not exist in the world. It makes sense for the perceptual system (from a computational point of view) to store merely a prototype memory of cats, instead of storing exemplar memories of each specific cat one has seen[50].

Simulated similarity priming takes it that perception associates its current representation with representations *generated* (synthesized) through some imagery procedure[51]. In this case, when one sees a cat, one primes in the manifold a *simulated* cat representation, one that is neither an exemplar memory trace nor a cat memory prototype. Crucially, such generative simulation ability should *not* be thought of as a conceptual ability. Rather, it is a purely *perceptual* ability: one can simulate similar representations to the occurrent one simply by exploiting a structured perceptual representation. Thus, one might slightly modify the value of one or several dimensions of the occurrent structured representation, generating a new representation. Because this representation remains very similar to the occurrent one, it is expected that it should be in the same object manifold. For instance, imagine I have never seen a golden chanterelle mushroom before. When I perceive one for the first time, I can perceptually simulate *other* golden chanterelles: I can simulate an object with a slightly darker skin, a slightly more angular shape, a slightly

[50] Exemplar and prototype storage are not necessarily opposed, but they might be viewed as complementary. We might store specific *exemplar* memory traces of particulars we have seen recently. But through time and necessary memory degradation, those exemplar traces might be lost, and the only memory trace we preserve is some kind of prototype memory. For instance, you might store an exemplar perceptual memory of the cat you saw this morning (so that you can instantly recognize it when re-encountering it in the evening). But several years after, it is dubious that you have stored a specific exemplar memory trace of this cat.

[51] I prefer to talk of *simulation* instead of *imagination* because I wish to allow for such simulation to be purely sub-personal.

smaller size. This simulated representation could still belong to the same golden chanterelle object manifold.

Fascinatingly, this generative capacity has been empirically demonstrated in a study by Henning Tiedemann and collaborators (Tiedemann et al., 2022). In this study, they showed that when presented with pictures of artificial objects (objects not classifiable a priori in a known category), participants can easily simulate and draw new exemplars of objects that they think could pertain to the same category. When shown to other participants, these drawings are usually classified together well above chance, showing that these are seen as pertaining to the same category. And the authors of the study conclude, in a very Kantian spirit:

> Taken together, the results of our experiments suggest that humans are not merely passive observers, who assign objects to categories based on their relative similarities in a fixed feature space. Instead, a key aspect of human generalization is our ability to identify important signatures of generative processes and then to run internal routines that actively synthesize new objects that we have never actually observed. (Tiedemann et al., 2022, p. 23)

The idea of simulated similarity priming is fascinating because it shows that we could classify an object even after one unique encounter with it (a process authors call "one-shot generalization"). Upon seeing a golden chanterelle for the first time, I might thus see it as being similar to *possible* simulated golden chanterelles. Just like with exemplar and prototype similarity priming then, with simulated similarity priming we do not perceive particulars in isolation, but in relation with similar representations in the same perceptual manifold. In the case of *simulated* similarity, these are *possible* representations of objects, not strictly speaking objectual memory traces.

Similarity priming that is purely simulated through internal routines must extremely rare though. It only occurs in cases when one has never seen an object like the one currently seen, This situation probably only occurs when subjects are presented with artificial stimuli, such as those created in psychology labs (as in Tiedemann et al., 2022), or with unconventional objects such as those stemming from the fertile imagination of a painter. But in almost all situations, one will have *some* memory traces that are similar to what one is seeing. Even if one does not prime specific memories of golden chanterelle exemplars or prototypes when seeing one (because one *doesn't* have such specific memories, or doesn't structure in a relevant way), one will still be able to prime some memories of *mushrooms*, or more generally of *plants*. Of course, after having been familiarized with golden chanterelle, one will be able to prime specific memories of golden chanterelle, instead of memories of mushrooms in general. This latter case is a case where one learns to perceptually structure in a more fine-grained way, i.e., to create submanifolds within the mushroom manifold.

While *pure* simulated similarity priming is rare, I think though that some simulated elements enter into most, if not all, kinds of similarity priming. There are in fact good

reasons for doubting that exemplar, prototype, and purely simulated similarity primings can be psychologically distinguished very neatly. "Generationist" or "constructivist" theories of memory (Michaelian, 2011), recognize the general idea that memory "produces content in addition to that which it took as input" (*ibid*., p. 324, see also De Brigard, 2014). In our case, this might imply that the representations primed are not pure "exemplars", "prototypes", or "simulations". Rather, they might be composites containing as parts some memory traces and some simulated complements. I think prototypes are such kind of composite representations: they are like simulated representations in the sense that they were never perceived before; they are like exemplar memories in the sense that they are summaries of them. When I talk of similarity priming as being a kind of *memory* priming, one should remember that this memory priming is actually not *purely mnemonic* but might often integrate simulated elements.

In any case, what is important for us in the ambit of the admissible contents of perception debate is the idea that *structured* perceptual representations can systematically *prime* similar perceptual memory representations. I call this process *schematization*, in a winding homage to Kant's schematism (though it should be clear that my own psychological understanding of schematization is quite different from that of Kant).

Because the structured representation has lost many of its feature representations (e.g., that of orientation, that of illumination etc.), the similar representation retrieved from perceptual memory might be similar to the current percept only in respects to some a very limited number of dimensions. For instance, a cat percept might automatically make us retrieve perceptual memories of cats, even if the currently seen cat is seen in an orientation or from a viewpoint that does *not* correspond to any perceptual memory[52]. Furthermore, one should keep in mind that the relevant dimensions on which similarity priming is based to not have to be features of whole objects (even if it can). Importantly, mere object *parts* might be sufficient to prime relevant perceptual memories in the same manifold as the entire object (e.g., the tail of a cat primes memories of cats, the wing of a plane primes memories of plane). It is empirically demonstrable that we can recognize objects merely by seeing some limited parts of them, and that these object parts elicit object-specific activity in higher-level visual cortex (Holzinger et al., 2019). This capacity makes sense in a environment where objects are most of the time perceived in occlusion.

Crucially, schematization is to be distinguished from categorization. Schematization does *not* make use of concepts stored in semantic memory. Rather, schematization only associates a current structured perceptual representation with perceptual memories through similarity priming. Now, as we already hinted at (see **Figure 6**), schematization can be

[52] Because the similarity of a current percept seen from an unusual viewpoint (e.g., a cat seen from above) to some perceptual memories that store more usual, canonical viewpoint requires more structuring (in this case, requires structuring by filtering out orientation), it might be that schematization is in this case more difficult and slower (Palmer et al., 1981). It is easier to schematize a cat percept (i.e., to retrieve relevant perceptual cat-like memories) when it is seen from a canonical viewpoint than a non-canonical one.

causally related to categorization. One can retrieve specific concepts associated with the primed representations and apply them to the current percept. This is what similarity-based views of categorization argue for (unsurprisingly, similarity-based views of categorization divide depending on whether the retrieved representations are thought to be exemplars or prototypes, as we saw in III.2.2 PSYCHOLOGICAL MODELS).

Schematization only appeals to *perceptual memories* (or perceptual simulations), while categorization requires access to semantic stores of concepts. This distinction between schematization and categorization is crucial psychologically, even though it might be hard to distinguish them at a personal, conscious level. In fact, schematization might just automatically trigger the retrieval of some concepts corresponding to the primed similar representations, i.e., automatically trigger categorization. But schematization and categorization do not yield the same kind of representations. Schematization is a kind of purely nonconceptual classification. It allows to recognize objects without categorizing them.

The distinction between nonconceptual schematization and conceptual categorization is further warranted by the possible observation of psychological dissociations between them. One can schematize while not being able to categorize. We can recognize someone but be unable to think about who that person is. Recognition only requires schematization. Categorization requires concept application. Infants and nonhuman animals are probably able to schematize (and thus to recognize), even though they might not be able to exploit such schematization in order to categorize, because they might not possess concepts to apply in the first place. A baby recognizing her mother's face is a plausible example of such schematization without categorization. The baby is able to (reassuringly) see her mother's face as being similar to something she has seen before (her mother's face). This makes her smile. To account for this behavior, one does not need to argue for the idea that the baby in some way applies a core concept of MOTHER'S FACE to what she perceives. And one doesn't need either to fall back to some kind of behavioristic interpretation whereby the baby's smiling is interpreted as resulting from pure behavioral conditioning. Schematization stands as a more plausible middle-ground explanation.

A laboratory exemplification of the schematization-without-categorization idea is given by cases of *associative perceptual agnosia*, or the incapacity to categorize what one is seeing. A patient with associative perceptual agnosia might still be able to see something as familiar (see something as something else she has seen before) but be unable to exploit such familiarity in similarity-based categorization because of some cognitive impairments. It's been for instance observed that patients with *facial* associative agnosia ("prosopagnosia"), while unable to categorize faces as being that of specific persons such as their husband's or children's, still show signs that they are (unconsciously) recognizing the faces presented (e.g., as detected through skin electrodermal responses). In a classical demonstration of this phenomenon, Daniel Tranel and Antonio R. Damasio thus present their result as follows: "The positive electrodermal discrimination is interpreted as an index of the rich co-activation of records that is produced when familiar facial stimuli successfully trigger the

reassembly of previously acquired representations of the physical structure of faces" (Tranel & Damasio, 1988, p. 235). This, in my own framework, can be interpreted as schematization-without-categorization.

Oppositely, we might also easily conceive of categorization without schematization. This is especially true if we consider *rule-based* types of categorizations. If Anne tells Sarah that everybody in the seminar room with a blue badge works in the HR-department, then Sarah will be able to categorize any person with a blue badge as being a person working in the HR-department. This hybrid perceptual-cognitive categorization has nothing to do with schematization: Sarah is simply following a rule she has learned to apply some concept (WORKS IN THE HR DEPARTMENT) to people she is seeing[53].

Such categorization-without-schematization can be further empirically observed in perceptual agnosia. Agnosia might be explained as an incapacity to bound normally perceived parts into whole objects, i.e., as an impaired perceptual structuring capacity. Because of this impaired structuring capacity, one might in some cases be unable to prime relevantly similar objects, disturbing schematization, *even though semantic knowledge about objects remains intact*. Remember the agnosic patient described by Riddoch and Humphrey who couldn't recognize images of baby carriage, but only reported seeing "wheels with spokes". This agnosic patient is interpreted by Riddoch and Humphrey as having trouble integrating the parts of the object in a coherent baby carriage whole. This patient clearly has disturbed perceptual structuring capacities, and it is expected that he will schematize through irrelevant dimensions (the picture for him does *not* resemble that of a baby stroller but resembles merely that of a wheel). But agnosic patients are reported to be able to categorize objects when these objects are *verbally described* to them. So, for instance, the agnosic patient might be able to categorize the image as one of a baby carriage if he is told that it represents an object with the function of transporting babies. The agnosic might *think* 'this is a BABY CARRIAGE', even if he is unable to visually recognize it. Prosopagnosic patients have also been reported to be able to recognize faces through a very slow and cognitively taxing exercise of feature-based identification taking up to several tens of seconds (Marotta et al., 2002)[54]. In my own vocabulary, I think it is better to view such prosopagnosic patients as performing a (cognitively difficult) task of face-categorization, instead of *recognizing* (through schematization) faces.

I take it that schematization is the psychological mechanism underlying genuine *perceptual* recognition. I oppose views that consider that recognition depends on categorization (or concept application). I think one can recognize without categorizing. I think *re*-cognizing is fundamentally dependent on being sensitive to the similarity of what

[53] Though if she repeats her perceptual experiences with people from the HR department, she might begin to learn to schematize in a way that allows her to perceptually recognize people from HR without having to compute over any rule.

[54] Face classification is achieved through "a time-consuming analysis of individual parts when the stimuli do not activate the face-specific processors" (Marotta et al., 2002, p. 45), i.e., when the stimuli are not (holistically) structured relevantly.

we are perceiving to what was perceived before. Such sensitivity is orthogonal to conceptual capacities. It is a purely *nonconceptual* capacity.

One instant suspicion about the process of schematization that I want to answer straight away regards its potential phenomenological implausibility. After all, when we see something or someone, we do not seem to be recalling any kind of mnemonic traces corresponding to this object (like when we recall persons that were with us in middle school), nor do we seem to consciously imagine a prototype or simulated exemplar of this object (as when we imagine a pink elephant in a rollercoaster). Why should we believe that schematization happens then?

The answer is that schematization does not imply that one ever *consciously*, *explicitly,* or *declaratively* primes memory representations and consciously recalls them. Schematization can function merely by sub-personally matching a current percept with *unconscious*, *implicit*, *non-declaratively* primed representations. A face might look familiar in certain respects, but one might not be consciously reminiscing about the face to which it looks familiar in such respects. When we recognize someone, we do so because we have seen this person before, have stored memory traces of this encounter, and are able to match our current perception of this person with such memory trace. One need not be conscious of the memory traces to which the current perception is matched. One person might just be incapable of explicitly remembering to whom the person looks familiar, even if actively trying to recall it ("I saw him before, but I can't figure out who that is").

Schematization is also implicit in the following sense: when one sees an apple, perceptual memories (together with simulated elements) are non-intentionally, involuntarily, automatically produced. Similarity priming in schematization is simply the automatic activation of neighbor (similar) representations of the occurrent structured percept. Just like a stone thrown in water automatically activates its surroundings, a structured perceptual representation automatically activates and simulates similar memory representations.

A useful way to think about schematization is also under the classical model of *semantic* priming. In classical *semantic* priming, the representation of one concept (e.g., DOCTOR) automatically primes the processing of semantically related concepts (e.g., NURSE), as observed typically through lowered reaction times in lexical decision tasks. Crucially, such priming is typically *implicit*: one does not consciously think about these semantically related concepts. They are rather *unconsciously* primed, facilitating later conceptual tasks. Such semantic priming effects are expected within the "semantic pointer" view of concepts I advocate for. Each concept is expected to be associated with a whole network of other concepts, as well as with modally-specific representations retrieved from perceptual memory.

A parallel kind of implicit priming is at work in schematization. When one perceptually represents, this representation implicitly primes *similar* representations from perceptual memory. One does not consciously recall or visualize the perceptual representations that are primed. Schematization unfolds outside of our conscious awareness

and intentional control. Schematization is automatic (though how it unfolds can be *causally* dependent on beliefs and intentions, as we saw with how one can change similarity relations through attentional control). If I'm travelling in a foreign country, there can be a situation where I cannot help but see the waiter as looking like my brother, even if I know very well that it is *not* my brother.

While perceptual similarity priming as involved in schematization has common characteristic with semantic priming in its automatic and implicit nature, we should keep in mind important differences. Contrary to *semantic* priming in thought, *similarity* priming in perception is not governed by conceptually-sensitive association laws. When thinking of DOCTOR, one might prime the concept NURSE because it is a concept that is close to DOCTOR in one's conceptual network (or a concept that is pointed to by my concept of DOCTOR). Semantic priming is fully orthogonal to perceptual similarity: DOCTOR can also prime STETHOSCOPE, PILL, or DISEASE, which bear no perceptual similarity to perceptual representations linked with the concept DOCTOR.

Similarity priming, on the contrary, relies *purely* on associative laws relying on perceptual content. One automatically primes *perceptually* similar (memory or simulated) representations to the occurrent percept. If one sees a stethoscope under certain relevant respects (through perceptual structuring), one implicitly and automatically primes perceptual representations that are similar to the occurrent stethoscope (that are in the same stethoscope representational manifold). If the perceptual structuring is specific enough, then the primed representations should be one of stethoscopes. Such schematization might *then* ground the possibility of *categorizing* the object as a STETHOSCOPE. Of course, such categorization might then trigger a process of *semantic* priming; one sees a stethoscope, thinks of a STETHOSCOPE, and then semantically primes DOCTOR and HOSPITAL, which might itself be making one picture doctors and hospitals. But *similarity perceptual priming* from seeing a stethoscope shouldn't directly prime perceptual representations of doctors and hospitals since doctors and hospitals are not perceptually similar to stethoscopes.

Finally, I conjecture that perceptual structuring, similarity priming, and perceptual memory are tied together in a virtuous schematization circle. Repeatedly seeing the same person (Joe Biden) or class (cats) help consolidate a large pool of perceptual memory traces for such person or class. Each time I see a picture of Joe Biden, I implicitly re-activate (through structuring and similarity priming) perceptual memories corresponding to Joe Biden. Re-activation consolidates such memories and makes it more probable that one will automatically prime them next time one sees an occurrence of the individual or the class. Trivially, it is hard to recognize someone we haven't seen for a long time (especially if we haven't had many encounters with that person). Likewise, when my grandma takes me to find golden chanterelles in the forest, I might each year have lost perceptual memories of golden chanterelles, making me incapable of recognizing golden chanterelles again (since I'm unable to prime any similar memory of golden chanterelles when I see one). While I retain some golden chanterelles perceptual memories at the end of my holiday week with grandma (we go mushroom hunting every day!), I soon forget some memory traces since

never during my urban life do I have to recognize golden chanterelles (I never re-activate and thus consolidate my perceptual traces of golden chanterelles), My grandma, on the opposite, has *lots* of consolidate memory traces of golden chanterelles, and she (or rather, her visual system) easily primes them when encountering a golden chanterelle.

Furthermore, the more one perceives a specific individual or class, the more precise we can expect the perceptual structuring to be, and the more efficient we can expect similarity priming to be. Each time I see a new golden chanterelle, not only do I re-activate and consolidate my perceptual memories of golden chanterelles, but I thereby store a *new* exemplar memory of golden chanterelle (or updates its memory prototype). The perceptual memories I now store are structured: they have dimensions that have been representationally sensitized for relevant similarity priming. As I relentlessly train to recognize golden chanterelles with grandma, my perceptual structuring becomes much sharper and more precise. Through grandma's feedback, I have come to learn that the most important dimension to recognize golden chanterelles is that of texture. I thus learn to structure my representation of golden chanterelles around the texture dimension, and this makes me perceptually prime distinctive memories (i.e., only those with that precise texture), which happen to be only memories of golden chanterelles. To grandma's pride, I have come to acquire a new schematization skill by learning to structure in golden chanterelle specific way.

After having acquired such skill, the perceptual memories I store of golden chanterelles will be memories that retain only relevant information: since every time I see and recognize a golden chanterelle, I emphasize their texture dimension in perceptual structuring. Afterward, similarity priming will thus be made even more efficient since my perceptual memory traces of golden chanterelle are structured in the same way as to how I perceptually structure them. My perceptual manifold corresponding to golden chanterelles is shrunk and made even more coherent (similarity relations increase). Thus, the more I perceptually structure in specific ways, the more I store specifically structured memories, the more similarity priming is eased, and the more I am incentivized to continue perceptually structuring in specific ways. A virtuous schematization circle has been created, which might lead to developing some kind of schematization expertise.

I think my upholding schematization in perception, or the idea that perception is not only depended on the currently perceived object but is also infused with memory and simulation (imagination) through similarity priming, bears many echoes in the history of philosophy. Immanuel Kant and Henri Bergson are to me clear examples. Thus, for Bergson:

> [E]very attentive perception truly involves a reflexion, in the etymological sense of the word, that is to say the projection, outside ourselves, of an actively created image, identical with, or similar to, the object on which it comes to mould itself. (Bergson, 2010, p. 124)

And:

> "[C]omplete perception is only defined and distinguished by its coalescence with a memory-image, which we send forth to meet. […] Does not this amount to saying that distinct perception is brought about by two opposite currents, of which the one, centripetal, comes from the external object, and the other, centrifugal, has for its point of departure that which we term 'pure memory'" (Bergson, 2010, p. 162-163)[55].

I said that my use of the term "schematization" was a winding homage to Kant's Schematism in the *Critique of Pure Reason*. Interpreting Kant's philosophy of mind is a tricky matter, and I do not pretend to have any expertise in such interpretative issues. But I at least trust a much more authoritative source than me, i.e., Peter Strawson. Interpreting Kant's ideas about Schematism, and especially the place of Imagination, Strawson thus writes:

> We have seen that there would be no question of counting any transient perception as a perception of an enduring and distinct object unless we were prepared or ready to count some different perceptions as perceptions of one and the same enduring and distinct object. The thought of other actual or possible perceptions as related in this way to the present perception has thus a peculiarly intimate relation to our counting or taking – to our ability to count or take – this present perception as the perception of such an object. This is not of course to say that even when, for example, we perceive and recognize (reidentify as the object it is) a familiar particular object, there need occur anything which we would count as the experience of actually recalling any particular past perception of that object. (It is not in this way, either, that imagination comes into the picture.) Indeed the more familiar the object, the less likely any such experience is. *Still, in a way, we can say in such a case that the past perceptions are alive in the present perception*. (Strawson, 2008, p. 58)

I do not know, and do not care that much, whether Strawson's interpretation of Kant is correct or not. I just want to expose clear convergences with my own empirically-based defense of schematization. Like Strawson, and like vision psychologists, I do think that there is a kind of memory involved in most of our daily perceptions, even though it is present only implicitly. I take myself to be perpetuating these old ideas with new theoretical and empirical means. It will come about at several times that much of the empirical literature in vision perception and visual memory echo in remarkable ways these armchair psychological intuitions.

[55] The entirety of *Mater and Memory*'s Chapter 2 ("Of the Recognition of Images. Memory and the Brain") is dedicated to exposing this idea of a constant welding of perception with memory, which for Bergson should not be distinguished. Of course, I only agree with this general idea at a general level. I do not subscribe to Bergson's denial of memory traces in the brain, and I subscribe to distinguishing between different kinds of memory (semantic, episodic, perceptual). I also obviously do not subscribe to Bergson's proto-enactivism and anti-psychologism.

IV.2.3. The Perceptual Nature of Schematization

I have argued that *non-conceptuality* is a good criterion for determining whether a psychological mechanism is *perceptual*, since such criterion allows not to trivialize perceptual Liberalism. Schematization strictly respects the non-conceptuality criterion. It does not access conceptual contents stored in semantic memory (i.e., amodal semantic pointers which are associated to perceptual representations and to verbal information, such as information about typical features of an object). Rather, schematization relies solely on non-conceptual processes: perceptual structuring and similarity priming. Structuring is the building of a streamlined, simplified representation that is exploitable by further classification processes (including similarity priming and rule-based categorization). Structuring only computes over low-level perceptual representations. It can be considered parts and parcels of perception because we have no reason to think that it *constitutively* depends on cognition (though it might *causally* depend on it). Structuring is a natural feature of perceptual processing as described by hierarchical models from vision sciences and computational implementations.

Similarity priming is also a purely non-conceptual process. In fact, I have been attentive to separate between the kind of similarity one gets from matching the current perceptual representation with *perceptual* memory traces, from the kind of *conceptual* similarity one gets from applying the same concept (categorization) to two representations. A lesson that should be clear at this point is that similarity priming from a structured representation is *not* a form of categorization in the classical sense (even if categorization might rely on similarity priming). Similarity priming does not apply concept drawn from amodal semantic stores. It (implicitly, sub-personally) matches *perceptually* similar representations: the current perceptual representation with perceptual memory traces.

That schematization is perceptual is further vindicated by the fact that, at least prima facie, it plausibly fulfills common hallmarks for perception as described in II.5 Common Hallmarks Arguments. Schematization is plausibly a mechanism that is fast, stimulus-dependent, automatic, sensitive to adaptation effects, and works in parallel. I do not possess irrefutable empirical evidence for this idea. I think that the purely non-conceptual nature of schematization remains the most crucial argument for its perceptual nature. Noticing that schematization also plausibly bears classical hallmarks of perception only comes to abductively strengthen this conclusion.

Schematization is plausibly *fast* and *automatic*. Structuring can be hardwired and automatized in neural processes through phylogenetic selection or ontogenetic learning processes. Some low-level feature values might just automatically project the low-level representation into specific structured spaces. For instance, low-level representations of ovaloid mid-sized objects with a face-gestalt might automatically activate within a "face" subspace which is particularly sensitive to such ovaloid shape values and face-gestalt values. In that face subspace, face representations would be highly similar (they form a coherent manifold) since they share the features to which the space is particularly sensitive. Non-

face representations on the contrary would be strongly distinguished from the face manifold.

Neurally speaking, this could be implemented by a specialized neural patch (e.g., the fusiform face area) which is exclusively sensitive to such particular feature values. Once the FFA is activated, similar representations are easily co-activated (through memory reactivation or simulations). It is known empirically that the FFA activates extremely fast, producing the typical N170 ERP 170 ms after stimulus onset (Hadjikhani et al., 2009). Furthermore, it is also known since its discovery that the FFA responds automatically when relevant features are detected, be they genuine animal faces, or mere organization of low-level features organized like faces (yielding the well-known phenomenon of face pareidolia).

The fast and automatic nature of face detection is good evidence for the idea that face recognition is not a process of slow conceptual interpretation (after all, we just know that a cloud cannot be a face), but rather a perceptual process of schematization. Furthermore, at such high-speed, what we are describing is a purely bottom-up sweep of neural activity, an activation phase that cannot yet connect with top-down modulations from semantic hubs such as the ATL. Schematization can thus exhibit the properties of a fast and automatic bottom-up process, a typical hallmark of perception.

Such fast and automatic response has been reported for other specialized neural patches in the temporal lobe, such as the *parahippocampal place area,* which responds specifically to scene stimuli in approximately 80 ms, and to buildings in approximately 170 ms, while the *extrastriate body area* responds specifically to human body parts in 110 ms (Pitcher et al., 2012). These areas also quickly and automatically fire when their favorite features are observed, supporting again their perceptual nature.

While a lot of research has focused on these highly specialized neural areas in the temporal lobe for the recognition of faces, places, and bodies, perceptual recognition might also be grounded in a more distributed neural code for other kinds of objects (A. C. Connolly et al., 2012). This wouldn't threaten the idea of fast and automatic schematization. A structured psychological subspace might actually be implemented by highly distributed neural areas, corresponding to distributed neural sensitivity to various feature dimensions. Such distributed activity might also be quickly and automatically activated if each distributed patch responds to their own preferred feature in a synchronized way. Schematization in general is thus plausibly fast and automatic.

Schematization is obviously strongly stimulus-dependent. One does not schematize if one is not visually connected to some perceived particular and corresponding unstructured information registration of it to structure in the first place. Schematization is plausibly *subtly* and *unpredictably* stimulus-dependent. Some very small changes in stimulus values might dramatically alter the structuring process and the mnemonic traces that are primed. For instance, the very slight changes in temporal intervals between two object movements might cause one to see some stimulus as being similar to previous events that were all *causal* events, or all *non-causal* events, as is observed in the studies of Albert

Michotte and followers. This could be because the subspace corresponding to causal events is extremely sensitive only to a very fine-grained and limited ranges of temporal intervals between object movements (just like the FFA is extremely sensitive only to specific face-diagnostic features). What these structured representations prime are thus only similarly fine-grained spatiotemporal representations of causal events. Furthermore, it has recently been reported that activity in the ventral recognitional pathway can be strongly disrupted by small changes in picture features (Gaziv et al., 2023), a characteristic typical of subtly stimulus-dependent processing, and thus of perception.

Schematization might also plausibly yield adaptation effects. First, if low-level representations adapt, then the schematization process necessarily adapts too. For instance, if a specific color is what's emphasized by the structuring process, and if one visually adapts to this color, then we should expect that when presented with a neutral stimulus with a neutral color, the subject will represent (at least for a short period of time) the negative opposite of such color. In this case, the structuring mechanism for this stimulus will be disturbed, and one will structure in an opposite subspace (if it exists), priming some *opposite* mnemonic traces. This would be the case where we have *both* low-level adaptation and high-level adaptation (adaptation of the schematizing process).

High-level adaptations have been reported for some high-level contents such as face gender, face identity, face age, or face ethnicity (Webster et al., 2004). Adaptation for face stimuli is easily demonstrable because it is experimentally possible to create "neutral" faces by morphing face pictures from two extremes of a variation spectrum (e.g., morphing a typically male face with a typically female face to create a gender-neutral face). In these cases, adaptation effects were observed: subjects report seeing the neutral face as the opposite of the adapted face (e.g., seeing the neutral face as a male face when adapting to a female face). Adaptation to specific faces of individuals has been reported to depend on their similarity: one adapts to a particular face as defined by a level of image similarity (a face manifold), rather than to a specific individual (Xu et al., 2009). Adaptation still occurs if one is adapted to similar faces of different individuals. This provides good prima facie reason to believe that face recognition is independent from face categorization, and that face recognition depends on schematization. Since such face schematization can adapt, this is good further reason to believe it is a purely perceptual mechanism.

Finally, can schematization be processed in parallel, outside the scope of attention? I think so[56]. Subpersonal structuring and similarity priming still happen even when top-down conscious attention to a particular is absent. Even if one is not paying attention, one might still recognize Jack entering the room. In other words, one might still select features that are specific to Jack in the low-level visual representation even if that representation is not the center of conscious attentional focus. Once that structuring is completed, implicit similarity priming occurs, and one thus visually recognizes Jack.

[56] Even though I don't think this is a particularly *distinguishing* hallmark of perception, schematization would clearly *not* be perceptual if it was assimilated to some kind of attention-hungry, serial declarative thinking.

Schematization is compatible with empirical data showing above chance recognition even when target stimuli are not attended to, or are presented peripherally, or are presented at ultra-fast rates (Crouzet & Thorpe, 2011; Grill-Spector & Kanwisher, 2005). Of course, schematization for unattended stimuli might be pretty unprecise, since the low-level representation to structure in the first place might be informationally quite poor and unprecise themselves. It is thus quite probable that many kinds of schematizations are only possible if the target object is explicitly and consciously attended to, because the relevant feature for a specific individual or class might only be perceived if the particular is attended to. But there are cases where imprecise representations of some basic features are sufficient to recognize someone or something. For instance, it is known that *curvature level* is sufficient to recognize animals (Zachariou et al., 2018), or that only seeing small picture snippets of an object (e.g., a part of a plane's wing) might be sufficient to recognize it (Ullman et al., 2016). Curvature level or picture snippets might just be sufficient representations to feed into some perceptual structuring and to yield some kind of similarity priming. A represented curvature level in an image might prime representations of animals; a represented part of a plane's wing might prime representations of planes.

To wrap up, we have seen that it is at least prima facie plausible that schematization is produced fast, automatically, in a stimulus-dependent way, yields adaptation aftereffects, and processes in parallel outside of attentional focus. Neither structuring nor similarity priming go against this conclusion. This weighs in favor of schematization being a purely *perceptual* process. Of course, these appeals to supposed perceptual hallmarks for schematization is subject to the same kinds of objections and pitfalls that I have introduced in the second chapter (II.5 COMMON HALLMARKS ARGUMENTS). I thus wish to emphasize that the perceptual nature of schematization does not first derive from it possibly bearing such hallmarks. Rather, it is the *non-conceptual nature* of this psychological mechanism that first drives the conclusion that it is plausibly purely perceptual. The observation of presumed perceptual hallmarks for schematization just comes as a further abductive icing on the argumentative cake.

Schematization is thus a good candidate for being the uniquely perceptual mechanism we are looking for to establish high-level contents in perception, since it is a mechanism for *classification* (perceptual recognition) that, contrary to categorization, is a purely *non-conceptual* process. I think schematization ultimately provides us with a much more needed clear-cut criterion for classifying a perceptual content as low-level or high-level: high-level contents are those representations produced through schematization, while low-level contents are not (they come before schematization, and serve as inputs to it). Of course, I do not take it to be a premise of my argument (otherwise it would obviously be overly circular), and we will see in the next chapter that schematization yields representations that fulfil plausible independent criteria for having high-level contents unreducible to low-level ones. Before this though, I present some empirical and theoretical reasons to believe in the mental reality of schematization.

IV.3. Reasons to Believe in Schematization

IV.3.1. Experimental Evidence for Schematization

Schematization (structuring and similarity priming) is empirically plausible from many empirical observations stemming from neuroscientific studies and computational models of vision.

Schematization is supported by neurophysiological studies of the primate visual system. Through the method of *representational similarity analysis* (Kriegeskorte et al., 2008; Kriegeskorte & Kievit, 2013), neuroscientists can determine how *neurally* similar representations of objects are. The study of neural similarity is upheld as allowing us to determine the kind of information represented in the visual system. In fact, one can easily compare neural similarity to some behavioral measure of similarity. Behavioral similarity scores can be obtained through behavioral tests (e.g., through an odd-one-out task with pictures of objects where one has to decide which two out of three objects are most similar). Behavioral similarity reflects our *judgments* about how similar two objects are. Such judgments themselves are thought to reflect how we categorize objects: two objects are judged as more similar when they pertain to the same category. If, upon presentation of object pictures, neural similarity in the brain correlates with such similarity judgments, then we can conclude that that brain area represents in a category specific way[57]. In other words, representational similarity analysis can tell us whether some brain area organizes its representations into semantically coherent manifolds.

[57] Notice that this is different from saying that such brain area represents *categories* (or concepts). It might just be that such areas represent information that is then exploited by further categorization processes.

Representational similarity analyses demonstrate three fundamental results. First, they display a dramatic difference between similarity relations in early vision and in the higher level of the visual hierarchy (i.e., in the ventral temporal pathway), as illustrated in **Figure 7** below, left. Neural similarities in the early visual cortex are semantically anarchical: they do not reflect in any way the similarities provided by behavioral measures (e.g., two different pictures of apples that are behaviorally judged as being similar will not yield similar brain activity in early visual cortex). This shows that our similarity judgments do not reflect representation of low-level properties as represented in early vision. This shouldn't be surprising: early visual areas represent many fine-grained properties (e.g., localization properties, size properties, viewpoint-specific properties, illumination properties, orientation properties etc.), and such fine-grainedness upsets semantically-relevant similarity relations. Early visual cortex representations are too unstructured. They form a semantically tangled perceptual space in which similarity relations are semantically meaningless.

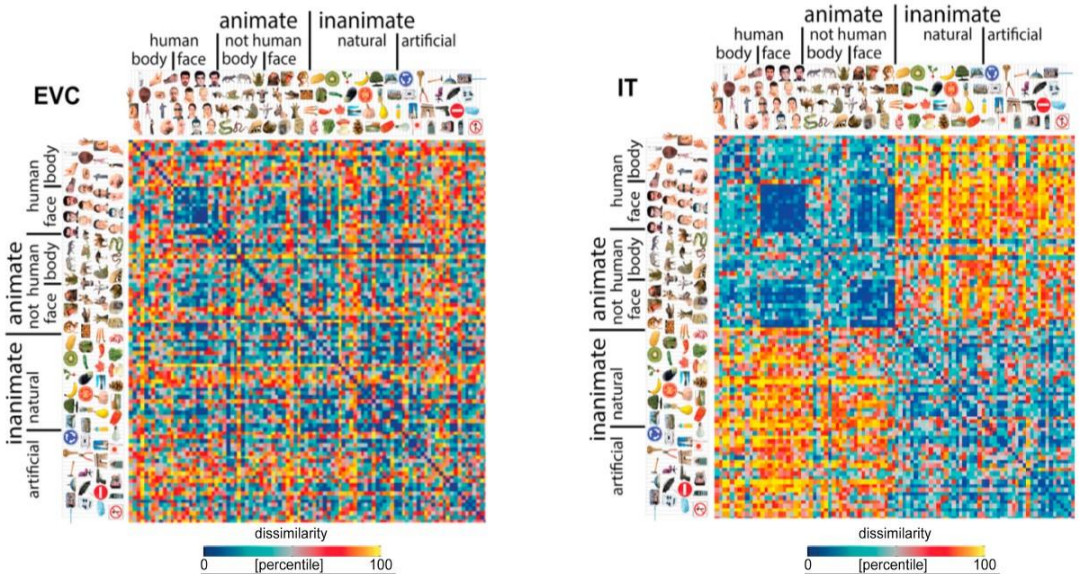

Figure 7. Representational similarity matrices obtained through fMRI data from early visual cortex (EVC) and inferotemporal cortex (IT).

The second important result from representational similarity analyses is that representations in more anterior visual areas of high-level vision (mostly, ventral temporal cortex) *do* have similarity relations that are semantically coherent (Figure 7, right). For instance, every

picture of faces will yield very similar neural activity. Similarity in neural activity is not limited to faces. Pictures of objects from similar categories tend to yield similar activity in the inferior temporal cortex (Huth et al., 2012)[58]. The fact that the brain can go from representations with anarchical, meaningless similarity relations in the early visual cortex to representations with meaningful similarity relations in the higher-level inferior temporal cortex (and in the parietal cortex) can be interpreted as the visual system structuring its representations. Here is how neuroscientists Nikolaus Kriegeskorte and Rogier Kievit interpret representational similarity analyses results:

> Overall, studies of the geometry of visual representations have impressively documented the stage-wise transformation of the retinal image from low-level representations of local physical features to high-level holistic representations of objects, faces, and scenes that better reflect perceptual qualities, *emphasizing behaviorally important categories and semantic dimensions and deemphasizing accidental and behaviorally irrelevant variation of the visual input* (Kriegeskorte & Kievit, 2013, p. 408, *my italics*)

What happens when one ascends from the early visual cortex to higher areas can thus be interpreted as a neural desensitization to certain categorically irrelevant perceptual dimensions, and a parallel emphasizing of categorically relevant dimensions. This is structuring neurally observed.

Now, we have seen that there exists an intense debate in the neuroscientific literature as to whether we should interpret neural activity in the inferior temporal cortex as being representationally selective for some categorically-diagnostic low-level features (feature-based coding) relevant for categorization, or whether it should be interpreted as directly implementing an abstract categorization (conceptualization) process (Bracci et al., 2017; Bracci & Op De Beeck, 2023). A third observation stemming from representational similarity analyses might favor former interpretation.

In fact, it has been observed that similarity relations in the inferior temporal cortex do *not* exactly reflect similarity judgments (**Figure 8** left, below). When making similarity judgments, it thus seems that subjects use information that is not available to the perceptual system. If the perceptual system were able to attribute the same concepts as we attribute in judgments, we could expect that similarity relations in high-level visual areas would parallel similarity judgments. But visual representations are irremediably feature-based, while judgments are not. Quite trivially, one cannot visually represent what is not visible. Thus, measures of similarity in the inferior temporal cortex of primates show that these remain quite insensitive to the category of "fruits", "plants", or "road signs". These are too diverse

[58] While some object categories yield similar activity because of the activation of highly specialized and localized brain areas (e.g., the fusiform face area for faces; the parahippocampal place area for scenes), most object categories yield similar activity because of some similar *population coding*, i.e., because of the same *distributed* brain activity in visual temporal cortex.

in terms of their perceivable features to cluster together in distinctive manifolds even after structuring processes.

Thus, results from the representational similarity literature describe a process whereby some unstructured perceptual representations are structured to keep only with some categorically-relevant feature dimensions that make representations similar in a semantically meaningful way. Such structuring is purely feedforward and automatic and does not depend on the application of concepts. It is not assimilable to some *judgment* of similarity.

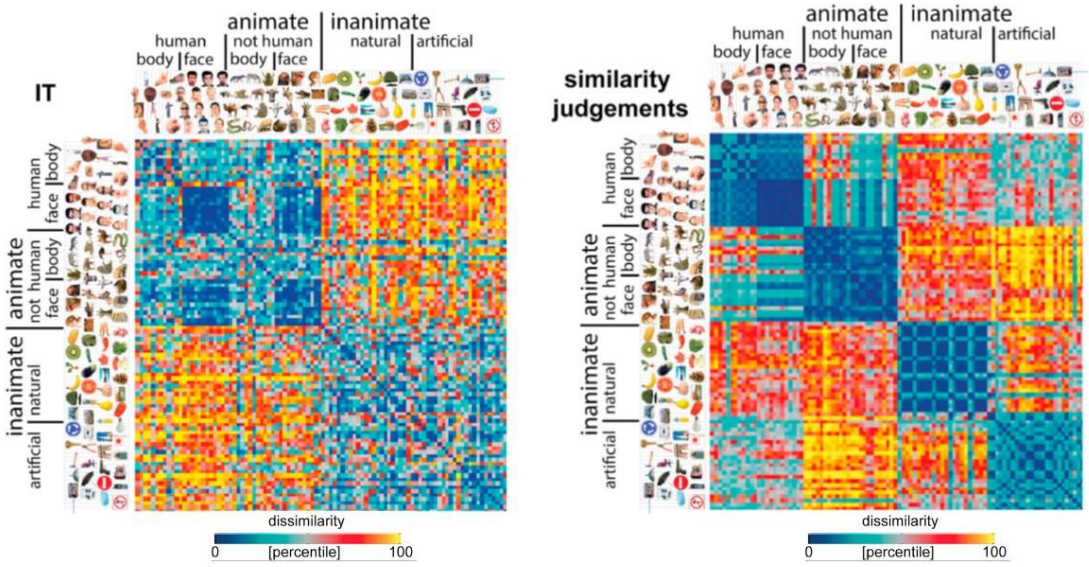

Figure 8. Representational similarity matrices obtained through fMRI data from inferotemporal cortex (IT) and similarity judgments.

Interestingly, it has been experimentally observed through magnetoencephalographic studies that such structuring process might not only be spatially organized along the visual hierarchy, but might also depend on *temporal* dynamics (Carlson et al., 2013; Contini et al., 2017). Thus, in the first hundred milliseconds after the presentation of a stimulus, representations in the ventral temporal pathway might be as unstructured as representations in the early visual cortex. Differentiation between different object representations is not possible at this stage. No meaningful similarity relations are observed. But after these first hundred milliseconds, object representations become much more discriminated (one and the same object yields similar neural activity). Structuring has begun, though similarity

relations do not yet cluster objects into meaningful manifolds. From 240 ms though, similarity manifolds distinguishing inanimate and animate objects are observed. Classificatory structuring is in place and remains at least for hundreds of milliseconds more. Interestingly, after some time, individual objects cease to be distinguishable within their manifolds, though the representational manifolds remain distinguished. This might indicate that object representations have lost most of their distinguishing features, in favor of representing merely category-relevant information. What is important to retain here is a clear demonstration of *temporal dynamic* structuring processes in perception. Such demonstration again emphasizes how much we should distrust our perceptual experience in accounting for the representational capacities of vision. Clearly, we do *not* have visual experiences that go from a nebulous slosh of unstructured low-level features, passing through discriminated objects, until purely categorical (e.g., positionless, colorless…) perception[59].

Strengthening such results from representational similarity analyses, one can also compare the similarity scores obtained from brain imagery to similarity scores obtained from the activity of deep neural network models (Zeman et al., 2020). The result of such comparisons shouldn't surprise us at this point: while the first layers of a deep neural network are extremely sensitive to localized information, and thus does not correlate with categorical information, the deepest layer of the network progressively lose such localized information and their representations become increasingly aligned with the categorically-relevant and filtered (untangled) representations in the visual temporal cortex. If one believes that such neural networks are a good model for understanding human visual processing, then this is yet again some good evidence for believing that representations along the visual ventral pathway follow a process of structuring.

Furthermore, one can additionally show through representational similarity analyses that representations of the last layer of deep neural networks do *not* satisfactorily predict how subjects name objects, i.e., do not predict subjects' categorical judgments (Devereux et al., 2018). To satisfactorily correlate with naming patterns, Barry J. Devereux and collaborators showed that one must add a *semantic network* at the top of the deep neural network. Semantic networks can represent non-perceptible information and manipulate linguistic-format information. This again shows that categorical *judgments* about pictures are not limited to purely visual information: categorical judgments result from the interpretation of filtered visual information by semantic systems. Schematization is not categorization, and this is demonstrated by how neural networks behave.

We thus have good neuroimaging and computational evidence that the visual system must learn to filter out some of its fine-grained (localized) representations to allow for perceptually-based generalizations, including categorical judgments. In other words, we have good evidence that the perceptual system structures its unstructured inputs.

[59] We only experience some intermediate step of the visual representational procedure (Prinz, 2012).

In such a structured perceptual space, if the perceptual system primes from memory some similar perceptual representation to the one that is currently perceived (a perceptual representation that is close by in representational space), then we can expect that such a retrieved representation will bear some semantic coherence with the occurrent percept. If one is currently seeing an apple, then the closest representation in the untangled perceptual space should be another apple.

We have good empirical reasons to think that the human perceptual system downstream, structured representations do produce such similarity primings. We already encountered some of these: people automatically recognize things they see, at a very high-level of precision (remember Brady's experiment in which participants are able to recognize thousands of pictures at a high level of accuracy). This indicates that, necessarily, some kind of mnemonic trace of the seen stimuli must have been stored and is retrieved when re-encountered again. Strikingly, such recognitional capacities have demonstrated to be potentially independent from capacities for explicit reminiscing through episodic memory. In one important demonstration, Joel L. Voss and collaborators thus demonstrated that, even when attention to a complex stimulus was strongly deterred during encoding (Voss et al., 2008), participants could still recognize it at later presentations with a high level of accuracy, *even if* they were unable to consciously recollect having perceived them.

This result corroborates a classical finding of the study of perceptual priming: priming occurs independently of explicit memory systems (episodic or semantic). We come back here to the idea of a purely perceptual, "presemantic", implicit mnemonic system as was described by Endel Tulving and Daniel L. Schacter (Tulving & Schacter, 1990). Their main thesis and evidence for it is worth emphasizing here, since it foreshadows my own view of similarity priming in perception (cf. III.1.2 SEMANTIC AND PERCEPTUAL MEMORY):

> First, the many different kinds of dissociations between priming and explicit memory can best be explained by the idea of a pre-semantic perceptual system that can operate independently of episodic memory. Second, neuropsychological studies have revealed dissociations between the reading of words and perceptual identification of other kinds of objects on the one hand, and semantic knowledge of words and objects on the other. These dissociations, too, point to a pre-semantic perceptual system that can operate independently of semantic memory. *We believe that priming and perceptual identification are of a single perceptual representation system (PRS), which exists separately from but interacts closely with other memory systems*. (Tulving & Schacter, 1990, p. 302; *my italics*)

Tulving and Shacter report dissociations between priming and explicit memory in many contexts: for amnesic patients, for very young infants (who have reduced capacities for explicit memory), for patients having ingested memory-upsetting drugs, and even for normal human adults (for whom we can experimentally manipulate explicit memory without affecting priming, and *vice versa*). For instance, amnesics perform better on a visual task (e.g., determining whether some picture of an object is physically possible) if they have

perceived the object before and had already determined if it was possible or not (Shimamura, 1986). This indicates that amnesics maintain some memory trace of the object, though they cannot *explicitly* remember the object[60].

Furthermore, perceptual priming effects are equally reported for agnosic patients than for non-agnosic subjects in non-semantic tasks, such as determining whether an object is a real object or not (Riddoch & Humphreys, 1987b; Schacter, 1990). This again indicates that similarity priming is a process distinguishable from the categorization of an object (which agnosic patients are by definition incapable of).

Words strongly prime phonemically and graphemically similar words: "nurse" perceptually primes "purse" (Hillinger, 1980). Again, this shows that there exists a memory system that only stores non-conceptualized, purely perceptual traces of perceived objects (including words), and which stored traces are automatically primed when similar objects are perceived (e.g., "nurse" priming "purse"). PRS does not contain abstract concepts, but concrete perceptual representations (either genuine exemplars or prototypes of them).

Tulving and Schacter also comment that the phenomenon of "reading without semantics" (the capacity to identify and read aloud words without comprehending them) is further evidence for an implicit mnemonic system detached from semantic memory altogether. In fact, identification of words "indicates that patients are able to gain access to stored representations of the words' visual form" (Tulving & Schacter, 1990, p. 304). This idea is especially compelling if we consider that such patients can correctly recognize *irregular* words, i.e., words which cannot be read following a "grapheme-to-phoneme conversion rule", such as "cough" or "blood". Reading might thus be a straightforward demonstration of similarity priming and schematization.

Tulving and Schacter note that similar priming effects are also observed with "nonverbal stimuli such as pictures, shapes, and faces". Interestingly, authors report that priming with 3D shapes only occurs if the subjects attend to the global structure of the shape. This is in line with the idea that much schematization depends on structuring the perceptual input's local spatial edge information into more general spatial gestalts or shapes. Mnemonic traces in PRS probably maintain such structured spatial information but discards much local information.

Similarity priming in schematization generalizes the point made by Tulving and Schacter about priming in lab-controlled tasks to everyday settings, an idea that is hinted at in the introduction of the paper: "It is reasonable to assume […] that priming represents a ubiquitous occurrence in everyday life" (Tulving & Schacter, 1990, p. 302). Perceptual priming advocates for a memory system that is *purely nonconceptual* and implicit. Similarity

[60] In my view, we could say that such amnesic patients have unconsciously recognized the object through (unconscious) schematization. Of course, they do not "recognize" the object in the classical sense of an explicit experience of recognition, but they "recognize" it in the sense that they unconsciously perceive the object as being similar to some things they have seen. Such unconscious recognition still has a measurable impact on their behavior, as demonstrated by their heightened

priming in schematization is neither assimilable to declarative episodic nor to semantic memory.

Schematization only assumes that such implicit perceptual priming is ubiquitous in everyday perception: when we see an object, the current perceptual representation automatically primes perceptual memory traces resembling it. Structured perceptual representations serve as *cues* for the priming of similar perceptual representations from such memory stores. Such mechanism is consensually observed in the lab, so there is no reason to doubt it shows in normal ecological perception.

Schematization, in order to be behaviorally relevant, needs to be a *long-term* priming effect. It's not that I see an apple as similar to other apples just in case I have seen apples some seconds ago. Rather, I can schematize something even if I have not seen an instance of that thing for several hours, days, or years (think of seeing a dear friend from college again). Perceptual memory traces must thus be convincingly long-term. This is not a problem. Studies of perceptual priming have shown that priming effects could be observed even at long time scales spanning up to several months and even years (Cave, 1997; Moutsopoulou et al., 2019; Schacter, 1990).

The idea of an implicit, presemantic "perceptual representational system" as classically described by Tulving and Schacter has been extended to a more general "implicit long-term memory" system, which is now not only postulated to account for priming effects, but also to account for other mental achievements such as visual category learning and object identification (Casale & Ashby, 2008; Cooke & Bear, 2015; Palmeri & Tarr, 2008). A crucial finding in this literature is that perceptual priming is not limited to mere *repetition* priming (priming for *exactly alike* repeated stimuli), but can be observed for stimuli that are sufficiently perceptually similar (Casale & Ashby, 2008; Cooper et al., 1992). For instance, Lynn A. Cooper and collaborators showed that priming effects are observed even after study-to-test transformations alter the size or reflection (left/right orientation) of the tested stimulus. This for the authors suggests that "a structural description system constructs representations of objects invariant over size and reflection, whereas a separable episodic system encodes these transformations as properties of an object's distinctive representation in memory" (Cooper et al., 1992, p. 43). This suggestion is not surprising under the schematization idea: perception structures its low-level representations so that they remain invariant (or less sensitive) to semantically irrelevant dimensions.

Still, perceptual memory still preserves a high-level of perceptual information as compared to semantic stores. Reviewing empirical research on visual memory, Thomas J. Palmeri and Michael J. Tarr thus conclude:

> Stable object perception is achieved by deploying our astonishing capacities for remembering particular experiences with particular objects under particular viewing conditions. We do encode a great deal of what we see as it originally appears. *Object perception is visual memory.* (Palmeri & Tarr, 2008, p. 172, *my itatlics*)

This astonishing memory capacity is proprietary to the perceptual system (see Burge, 2022, Chapter 17, for a related view), and it steadily infuses our daily perception of these objects. For Palmeri and Tarr, "[visual memories] for objects are part and parcel of the perception of those same objects, and object recognition is accomplished by comparing two perceptual representations" (Palmeri & Tarr, 2008, p. 166) and "memory representations of objects do indeed retain perceptual details […]. And, in fact, many of those detailed memory representations are the very representations that underlie visual object perception itself" (*ibid*., p. 167). This "Kantian-Bergsonian"-sounding theme of a perpetual concoction of perception and memory was also reflected in Tulving and Schacter's reflections on priming, when they advanced that "one remarkable feature of priming is that, unlike other forms of cognitive memory, it is nonconscious. *A person perceiving a familiar object is not aware that what is perceived is as much an expression of memory as it is of perception*" (Tulving & Schacter, 1990, p. 302, *my italics*).

The idea of perceptual similarity priming is further vindicated by the fact that the idea of a form of memory *within* perceptual areas has been made plausible by recent neuroimagining observations. The old strict modularism that distinguished between fully functionally distinguished brain areas (e.g., occipitotemporal and occipitoparietal streams for perception, and medial temporal lobe areas for memory) is now largely empirically outdated (Martin & Barense, 2023). It is in particular known that the perceptual streams themselves, down to the very earliest perceptual areas (Cooke & Bear, 2015) have memory capacities[61] that are independent from the activity of classical mnemonic areas in the medial temporal lobe (such as the hippocampus or the perirhinal cortex).

Precursor fMRI studies that started as soon as Schacter's and Tulving's theoretical description of perceptual priming relying on implicit perceptual representation system demonstrated that the areas implicated were distinguished from those involved in explicit-declarative memory (Reber et al., 1998). In particular, the areas involved in perceptual priming were found to overlap with those traditionally associated with perception (Wiggs & Martin, 1998). These latter researchers explicitly associate perceptual priming with a kind of *perceptual learning*, instead than with a kind of traditional explicit memory encoding, a theme which will take central place in VI KNOWING HOW TO Schematize.

The idea that experiences systematically leave some memory traces in the form of physical modifications is at least as old as Richard Semon's theory of *engrams* at the beginning of the 20th century. The basic idea behind engrams is that connected excitations in the brain produced by external stimulations should be bound together in some kind of unified complex that can be later reactivated. This engramic mechanism has now received extensive analysis in the recent neuroscientific literature (Josselyn & Tonegawa, 2020). A basic idea in this literature is that some specialized cells ("engram cells") undergo specific biochemical changes (e.g., synaptic plasticity) during the encoding phase to store

[61] It is also empirically demonstrated that the classical memory areas in medial temporal lobe perform some perceptual functions.

information. This information can then be retrieved by reactivating a particular group of engram cells.

Interestingly, contemporary theories suggest that such engrams are automatically formed in perception (instead of requiring explicit memorization in episodic memory). For instance, Hellen C. Barron and collaborators (Barron et al., 2017) have argued that, to restore equilibrium in excitatory perceptual activity, the brain automatically creates "inhibitory engrams" which function as inhibitory negatives of the actual pattern of excitation. The brain needs to *habituate* to stimuli to spare brain energy and to remain open to novel stimuli. Such inhibitory engrams are maintained and consolidated in the brain through classical processes of Hebbian plasticity. Such inhibitory engrams can later be "disinhibited", thus reproducing the excitation patterns that they aimed at inhibiting. Metaphorically, inhibitory engrams are really like negative films that are automatically produced in perception, negatives that are stored and can be "developed" at any time to yield the original picture. As Barron et al. conclude:

> In summary, inhibitory engrams provide a homeostatic mechanism that facilitates flexible yet stable storage. They greatly increase the storage capacity of neural networks, protecting against representational interference and runaway excitation. Furthermore, inhibitory engrams provide a gating mechanism that ensures information is stored silently and only released or expressed at time points that are relevant for behavior. (Barron et al., 2017, p. 6672)

Details aside here, the idea of inhibitory engrams as the neural mechanism underlying perceptual memory is perfectly compatible with my own view of schematization. We could imagine that, besides explicit, declarative recalls, what could activate a disinhibition process are simply occurrent perceptual cues which are similar to the inhibitory engram. This is just what we would expect from perceptual priming and recognition studies. Of course, this remains speculative and overly cursory. I do not aim here at defending that inhibitory engrams are necessary neural implementations for schematization. My point is simply that the idea of schematization lines up with many things we know about perception and memory.

IV.3.2. Schematization Avoids Overintellectualizing Behavior

Besides support stemming from neural, computational, and cognitive psychology, there is a general abductive argument to be made to support schematization. I think schematization should be regarded as our best explanation of many kinds of behavior, especially behaviors we want to explain without over-intellectualizing them, or supposing they rely on concept mastery.

One can start with behavioral observations about how we constantly exploit perceptual similarity in many crucial aspects of our lives. As we saw, there is massive

evidence in favor of at least some categorizations being similarity-based[62], and thus evidence for structuring and similarity priming. This evidence is particularly telling in the case of infants, who can exhibit sensitivity to categorically-relevant similarities between pictures (e.g., seeing two different pictures of cats as more similar than a picture of a cat and a picture of a dog) at such young ages as four months old (Hahn et al., 2010; Quinn et al., 2001). Since we can reasonably take it than infants do not possess corresponding concepts (such as CAT and DOG concepts), the best explanation for such sensitivity to similarity is in terms of schematization as described in the previous section: infants are able to represent that a current stimulus is perceptually similar to something they have seen before under certain respects.

It is even more telling in the case of simple animals like pigeons, who have also been reported as being behaviorally sensitive to similarities between different stimuli (Soto & Wasserman, 2014). Again, since we cannot reasonably presuppose that such simple animals like pigeons possess precise class concepts (they might just possess no concepts at all), a good explanation of their behavioral sensitivity to similarity relations is that pigeons achieve something akin to non-conceptual schematization: they recognize an object as similar to something they have seen before. The ethological literature on pigeon "categorization" (of course, in my framework, we should rather talk of pigeon *recognition*) converges on the idea that such achievement is allowed by a feedforward feature extraction process that relies exclusively on perceptual capacities, and not cognitive ones (Navarro et al., 2019).

Schematization explains many other kinds of behaviors that *generalize* from experience with some particulars to unknown but similar particulars. In conditioning experiments, "stimulus generalization" designates the fact that organisms can generalize their conditioned behavior to a specific stimulus to stimuli that resemble the conditioned stimulus (Mednick & Freedman, 1960). When an organism associates a certain behavior, affect, or thought with a specific stimulus (e.g., a dog associating food with hearing a certain bell sound), it can associate the same behavior, affect, or thought to similar stimuli (e.g., the dog associating food with a different bell sound).

That such stimulus generalizations occur is a good reason to believe that animals such as dogs schematize. The explanation we then get of stimulus generalization is simple: when the dog hears the bell, it automatically and implicitly primes similar memories of bell sounds it has heard before. The association of these primed memories with a specific event (food providing) is thus itself associated with the occurrent bell sound. This explanation is more attractive than an explanation that would require us to presuppose that the dog's behavioral generalization stems from a conceptualization of what it is hearing (BELL SOUND → food is arriving). This is the conclusion that old behaviorist psychologists working on conditioning have arrived at, though I obviously do not subscribe to the behavioristic

[62] Though I do not wish to argue that *all* categorizations are similarity-based. I think one can be a pluralist about categorization and accept that humans use both similarity-based and rule-based categorizations strategies.

conclusion that generalization is explained by non-mental stimulus-response automatisms. Rather, generalization is explained by a mental, representational (though non-conceptual) mechanism: schematization.

As we saw considering "Goodman's challenge", one cannot represent relevant similarity relations without determining some respects for similarity assessments. Determining respects for similarity might be done through the cognitive representation and manipulation of some rules: I can voluntarily only consider the texture of a stimulus because I *know* (store a rule) that 'texture distinguishes between category X and Y'. But we cannot expect all classifications to rely on such cognitive processes (especially for infants and nonhuman animals). What we should expect is a *non-conceptual* mechanism that selects dimensions for similarity comparisons. As I argued in the previous section, schematization appears as a good explanation for such non-conceptual classification capacities.

One could point out that we do not systematically need to appeal to schematization to explain classification or behavioral generalizations. Very simple organisms, living in extremely limited environmental niches and with a fixed set of behavioral generalizations, will probably have a very simple perceptual space that does not need to be schematized to yield generalizations. Such a simple organism simply needs to associate some fixed behavior to some fixed range of values in its perceptual space. For instance, a frog might just automatically snap its tongue at any small black moving object, without having to recognize it as similar in certain respects to things it has seen and snapped at before.

However, and crucially, as soon as the behavioral repertoire of an organism becomes complex enough, relying on primitive associations will not suffice anymore, and schematization will be called for. Especially, as soon as an organism needs *flexibility* in its generalizations, then relying merely on primitive fixed associations becomes impossible. Complex behaviors in complex environments require the organism to be able to represent many perceptual dimensions of particulars, and to be able to flexibly circumscribe *relevant* associations to context-specific and task-specific dimensions. This fact is quite blatant for organisms like humans, whose behavior is extremely flexible and cannot simply depend on primitive associations. Color is irrelevant to recognize a cellphone, but it is fundamental to recognize food. One needs to be able to separate the wheat from the chaff of which perceptual dimensions are relevant and which are not in specific situations and for specific goals.

Schematization just provides such flexibility: one can differentially structure perceptual representations depending on specific context and task demands. One can *learn* to schematize differently and can voluntarily modulate her structuring process through attentional modulations. Schematization thus grants us flexibility in the kind of similarity to which we are sensitive to. Schematization explains flexible context-sensitive behavioral generalizations without having to postulate overly intellectual capacities such as conceptual reasoning. The fact that schematization allows one to explain complex (generalizing) behaviors without having to succumb to "intellectualist" hypotheses (e.g., postulating that infants and animals have conceptual mastery) constitutes forceful abductive evidence for it.

The kind of generalization one can get from schematizing is not limited to stimulus generalizations as in classical conditioning experiments. Other important behavioral generalizations include all kinds of intentional actions. Again, the fact that human adults, human infants, and nonhuman animals perform many kinds of relevant intentional actions with never-seen-before objects seems to weigh in favor of schematization. Apes, infants, and adults can spontaneously represent that an apple is edible, even if they have never seen that particular apple before. This is because they automatically represent (schematize) this apple as being similar to other things they have seen before that were edible. Again, not only is schematization a potent explanation of this generalization capacity, but it is the *best* explanation because it doesn't presuppose any capacity for conceptualizing what one is seeing (it doesn't presuppose apes, infants, and adults see that something 'is an APPLE'). Moreover, schematization avoids the problems faced by classical associationist theories by being able to naturally account for generalizations with completely novel stimuli. Schematization keeps up with the ambition of explaining such generalizations without appealing to overly intellectualistic suppositions.

Other important behavioral generalizations that schematization can potentially successfully account for are word acquisition behaviors. Infants quickly learn to point at objects and express words about them. An infant can thus point at a cat and say "cat!", and she can do so pointing at a cat she had never seen before. Again, schematization explains this, and explains it in satisfactorily non-intellectualist way. Here, the fact that schematization does not presuppose that children need possess the concept CAT to say "cat!" when they see a cat is not only a desirable component, but it might simply be a *necessary* one. In fact, it seems that to account for the phenomenon of concept learning in children, we need to begin with some *non-conceptual* representations, which serve as the scaffold to arrive at conceptual representations. Concepts (semantic pointers) need to be able to point to some information in the first place. If we do not prioritize nonconceptual contents over conceptual ones, it seems one will have to assume a form of strong nativism about many concepts, which to many is not desirable. With schematization, an infant can begin with non-conceptual similarity grouping of particulars (object recognition) and associate a specific word to such similar particulars.

At first, a word is only associated with some perceptual schematization. The child's concept (e.g., the concept of CAT), is thus purely observational. After some time though, the child learns more information about the thing she calls "cats", and especially more non-perceptual information through language acquisition (she learns that 'cats are feline', 'cats were worshipped in Egypt'…). She thus comes to de-observationalize her cat concept, transforming it into a general non-observational concept. This concept is both pointing to perceptual information (a provided by schematization) and to verbal information. Mastery of the concept CAT depends crucially on being able to fix it (to *address* the pointer) to a particular set of representations in the first place (to bootstrap learning). This is allowed by schematization.

Schematization allows bootstrapping the acquisition of concepts (and not just explain the *application* of concepts). After all, we learn many of our concepts not only by being *told* about them, but simply by perceiving some of their instances. It seems we do not even need to be told *anything* to learn basic concepts: Daddy can simply point to the cat and say "cat!". Baby, being able to schematize, will soon correctly say "cat!" when seeing a new instance of a cat she has never seen before. Through development, baby learns to internalize "cat!" sayings in inner speech, and she develops a capacity for demonstratively thinking through internalized speech. The internalized "cat" saying might now stand as an autonomous representational vehicle in demonstrative thought that points to occurrent schematizations. This is the dawning of possessing a (purely observational) CAT concept.

Surely, schematization won't explain *all* kinds of concept learning. In fact, it seems we can master concepts even if we have never seen any occurrence of the things or things to which the concept refers. An Incan denizen in 1532 might learn from a conquistador many things about apples: that they are spherical, round, smooth fruits that range from green to red, that they come from another continent, that they yield one harvest at the end of the summer etc. I think we can attribute mastery of the APPLE concept to this Incan denizen. Notice that the Incan denizen might categorize something as an apple when he sees one for the first time through some kind of rule-based cognitive inference. But he might not yet be able to *perceptually schematize* it in apple-specific ways, since the first apple he sees does not look like anything he has seen before. He cannot *recognize* an apple.

The general argumentative point I make is that schematization can explain many kinds of complex behavioral generalizations: stimulus generalization in conditioning, perceptually-guided actions and habits, primitive word utterances and thoughts, similarity-based categorizations, *without having to over-intellectualize the possibility of such behaviors (without having to presume concept possession)*. This I think constitutes a good *abductive* argument for schematization, since such behavioral generalizations *are* observed, and since we *do* want to avoid over-intellectualizing them, especially so as to attribute them to infants and animals.

What I wish I have made potent is that the idea of structuring and similarity priming in perception are well-established empirical facts. Behavioral, neural, and computational models research support their plausibility. We have started out this chapter with what I called "Goodman's challenge": how is it possible to represent a semantically interesting similarity between some particulars without subsuming such particulars under the same concepts? We now have a neat solution that is empirically plausible: *schematization*. We can represent some particular as similar to some others through *implicit perceptual priming*. Such perceptual priming is semantically coherent because it is driven by representations that are *structured*: what representation is primed does not depend on just any perceptual feature, but on features that have precisely been structured and filtered so as to serve useful behavioral generalizations. Useful behavioral generalizations depend on being attuned to not just any arbitrary similarity relation, but to similarity relations which track relevant kinds in the world.

Schematization thus stands as a primitive nonconceptual mechanism for creating some equivalence between particulars. It is a purely perceptual mechanism of classification. It allows us to perceptually represent some particular *o* as equivalent to some memory representations M_i under some particular respects. Schematization stands as our best candidate for a perceptual mechanism grounding high-level contents in perception. In fact, the kind of content produced my such representational capacity isn't like any of the contents traditionally accepted by Conservatives. But it's also quite unlike the abstract contents proposed by classical Liberal views, which tend to be more like concepts (e.g., natural kinds or functional kind contents). What exactly is the semantic nature of the high-level contents supported by schematization? This issue is the topic of the next chapter.

V. ASPECTS

When we see an object, we visually represent some of its low-level properties: its shape, its color, its texture, its distance, its size, its position etc. Through *schematization*, we represent this object as being similar to other perceptual representations (stored in implicit perceptual memory). Let's take our usual apple example. When we see an apple, we see some spherical red 3D object with a concavity at the top and a small stem sticking out of it. These are all low-level features that are consensually recognized as being perceivable. At this point, we cannot say that we perceive something as an *apple*.

Through perceptual structuring, some of the low-level attributes will be discarded or deemphasized, and a sparsified perceptual representation will be produced in high-level vision. For the sake of simplicity, we can for instance imagine that only some determinate *shape* attribute of the representation is preserved: shape_{42} (the subscript indicates that we are talking of the determinate shape of the seen apple). Now, the representationally closest representations to shape_{42} are other representations of objects stored in perceptual memory with similar shapes: shape_{40}, shape_{41}, shape_{43}. Upon structuring the apple as an object with shape_{42}, one automatically and unconsciously primes these similar representations. One thus represents *o* as being similar to some other objects with respect to shape; or one represents *o* as shape_{42}, similar to shape_{41}, similar to shape_{40}, similar to shape_{43}. Saying that one perceptually represents *similarity* as such is misleading. Rather one *associates* mnemonic representations to the current percept through principles of perceptual similarity. So it might be more correct to write that one represents *o* as shape_{42}, associated to shape_{41}, associated to shape_{40}, associated to shape_{43}

I submit in this chapter that such associative representation produced by schematization constitutes a kind of high-level content in perception. I will call such representation as constituted by schematization an *aspect*. I will adopt a new typographical convention for aspect representation: I will often (though not systematically for reasons of textual fluency) use icons when writing about them. For instance, instead of writing that

one sees *o* as shape$_{42}$, similar-to-shape$_{40}$, similar-to-shape$_{41}$, similar-to-shape$_{43}$, we can say that one sees *o* as 🍎, as illustrated in **Figure 9**.

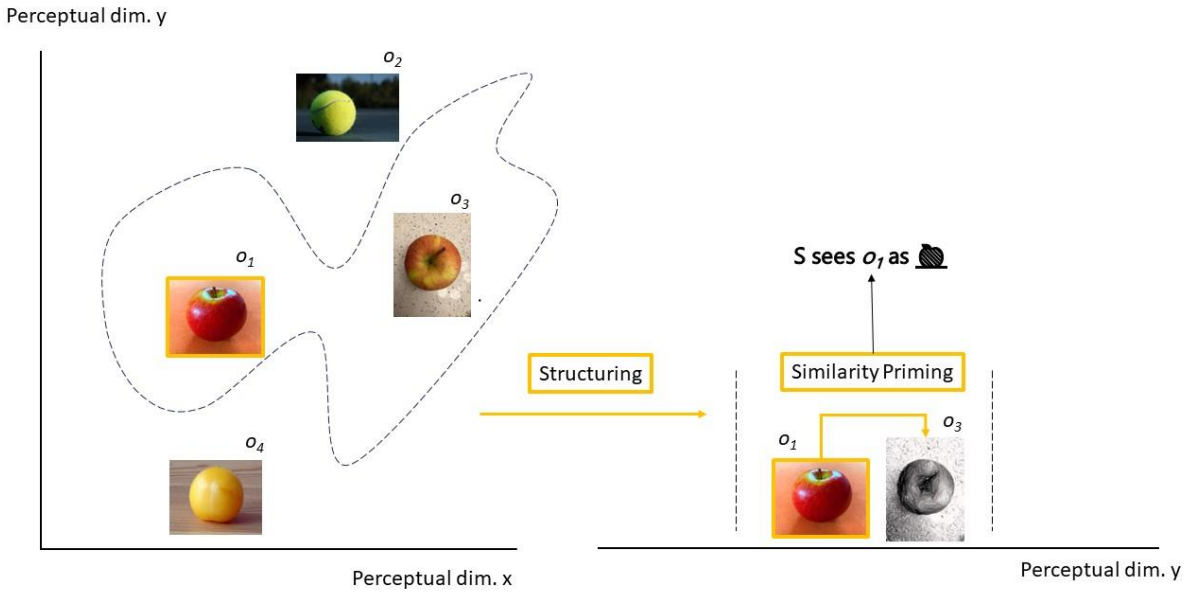

Figure 9. Schematization yielding an apple-aspect representation.

The goal of this chapter is to develop a better understanding of aspects, and especially of their semantic properties. A main theme of the chapter will be that aspects are proprietary high-level *nonconceptual* representations of perception. They are neither reducible to low-level representations nor to concepts. Their semantic is unique, in that aspects represent a special property of the world: *superficial kind properties*. A perceptual Liberalism based on aspects is thus a *moderate* Liberalism, one that accepts some high-level contents in perception though not the same kinds of contents as those included in thought. I will first (V.1) present the representational specificities of aspects (and especially their nonconceptual nature) by considering how they can be intuited from classical cases of perceptual bistability (thus paying homage to their Wittgensteinian origins). I will then argue that aspects are not reducible to low-level contents, and that they are perceptually proprietary *high-level* contents (V.2). Finally, I will show that such high-level aspect representations represent a specific kind content that semantically distinguishes them from the more traditional natural kind or functional kind contents of thought (V.3).

V.1. WITTGENSTEINIAN VARIATIONS

V.1.1. THE NONCONCEPTUAL REPRESENTATION OF ASPECTS

An apple aspect 🍎 represents an object as being perceptually similar to other objects under certain low-level (perceptual) respects. It is to be distinguished from the APPLE general concept, which involves a mastery of some minimal information about apples, both perceivable (apples are round, apples taste sweet, apples are smooth…) and non-perceivable (apples are the fruit of *Malus domestica* trees, their seeds contain amygdalin chemicals…).

Aspects are also to be distinguished from related *observational* concepts. Observational concepts are still amodal semantic pointers (just like non-observational concepts), though they only point to perceptual information. One might token a purely observational concept of APPLE when the information one tokens is only about what apples look like. Still, we talk of a *concept* in that we are dealing with a representation that can be tokened in the absence of a perceived object in propositional compositions ("she wants APPLES as wallpaper motifs"). Also, it is a matter of controversy whether the perceptual information pointed to by observational concepts is re-activated or "simulated" in some perceptual format (Barsalou, 1999; Prinz, 2002), or whether they point to other observational concepts (e.g., APPLE points to RED, ROUND, SOUR, SWEET), or a mix of both. Aspects shouldn't be assimilated to non-demonstrative simulation devices or to semantic pointers.

An aspect is a modally-specific, demonstrative perceptual representation. Aspects though are not reducible to *demonstrative* concepts. Besides my skepticism about the cogency of such representational entities (see III.1.1 THE NON-CONCEPTUALITY CRITERION), I think aspects remain different from demonstrative concepts because aspects are perceptual representations that are to be distinguished from propositional thought capacities (including demonstrative thought) and are grounded in evolutionary and ontogenetically primitive representational capacities. One cannot ground a cogent understanding of demonstrative thought (THAT(apple)), if one doesn't accept the existence of some pre-demonstrative contents to be demonstrated in the first place (J. Levine, 2010). Maybe a cogent understanding of the demonstrative concept THAT(apple) is as THAT(🍎). I will leave such issue aside here since I am suspicious of the relevance of demonstrative concepts anyway.

The use of the term "aspect" is intended as a direct hint to Wittgenstein's description of aspect perception as a way to account for bistability in ambiguous figures, such as the famous duck-rabbit image (Wittgenstein et al., 2009). Wittgenstein described what happens when one switches between seeing a duck a seeing a rabbit in the duck-rabbit

figure as switching between seeing the drawing *under different aspects* or seeing the drawing *as* different things[63].

Wittgenstein wrote that "the lightning up of an aspect seems half visual experience, half thought" (Wittgenstein et al., 2009, p. 207). It is pretty clear from his writing that he took the perception of aspects to go beyond and above the perception of simple low-level properties. The duck-rabbit drawing does not change while we stare at it, and it is a plausible idea to argue that the low-level properties we represent do not change either. It is only the aspect under which we visually experience the drawing that changes. For Wittgenstein, "[t]he expression of a change of aspect is an expression of a new perception and, at the same time, an expression of an unchanged perception" (Wittgenstein et al., 2009, p. 206).

Wittgenstein took aspect perception to be a motley phenomenon, subsuming the perception of bistable figures, but also perceiving a face as resembling another, seeing a picture of a horse as that of a horse *galloping*, seeing something as foreground or background, or seeing hidden figures in puzzle-pictures (as when we see a Dalmatian dog in "Mooney" images). Regarding the latter case of aspect perception, Wittgenstein writes:

> I suddenly see the solution of a puzzle-picture. Where there were previously branches, now there is a human figure. My visual impression has changed, and now I recognize that it has not only shape and colour, but also a quite particular 'organization'. —— My visual impression has changed – what was it like before; what is it like now? —— If I represent it by means of an exact copy – and isn't that a good representation of it? – no change shows up. (Wittgenstein et al., 2009, p. 206)

This idea that making a copy of what one sees cannot make aspects explicit is interesting because it shows that we are not dealing with low-level perception. This theme actually recurs in some contemporary arguments within the admissible contents of perception debate. Fred Dretske for instance has argued in favor of a Conservative position precisely by pointing out that a Liberal position should predict that someone who perceptually represents a high-level content (e.g., seeing a pine-tree) should make a different copy than someone who doesn't (Dretske, 2015). But we cannot conceive what would be different in such drawings, so Liberalism must be false. Of course, Dretske's argument can only take off if we accept the premise that everything that is perceptually experienced is drawable, a premise that might obviously be resisted (and one that Wittgenstein rejects).

---

[63] Wittgenstein reserved the term "seeing as" for specific cases of aspect perception. His use is thus more restricted than mine. I take seeing as to be constitutional of what seeing is. We always and necessarily see particulars as having certain attributes. Wittgenstein on the opposite took seeing-as as a problematic phenomenon that needed explanation. He also emphasized that he thought of aspect perception as a dynamic phenomenon. Aspects are described as "lighting up" (*Aufleuchten*). I do not emphasize this dynamic aspect of aspect perception. I think one might see aspects just as instantly and naturally as one sees shape and color. But I do concur with Wittgenstein in thinking that the aspect we represent a particular as having can change in time.

There is a large secondary literature about how exactly one should interpret Wittgenstein's somewhat sibylline remarks about aspect perception (Baz, 2020). It's been for instance harshly debated whether aspect perception depends on or is constituted by concept possession and application. I won't dwell on these exegetical disputes here. The approach to aspect perception that I advocate is clearly nonconceptual: aspect perception is the result of perceptual schematization. Aspect switches between the duck and the rabbit correspond to an unstable schematization: one can perceptually structure differently the *same* picture, by voluntarily attending to different parts of it, which result in different similarity primings (of duck-like perceptual memories or of rabbit-like perceptual memories). Thus, one switches from seeing the *same* picture as  and as  , which is *not* equivalent to seeing that the picture is a DUCK or is a RABBIT (Gauker, 2017; Orlandi, 2011).

This view of aspect switching understood as switchable schematization, thus switchable similarity relations, remains faithful to a fundamental intuition of Wittgenstein's:

> The colour in the visual impression corresponds to the colour of the object (this blotting paper looks pink to me, and is pink) — the shape in the visual impression to the shape of the object (it looks rectangular to me, and is rectangular) — but what I perceive in the lighting up of an aspect is not a property of the object, *but an internal relation between it and other objects*. (L. Wittgenstein et al., 2009, p. 223, *my italics*)

I gladly take on this intuition in my own framework: aspect representation is unlike color or shape representation in that aspect representation is relational and depends on *other* associated representations (i.e., implicitly tokened perceptual memories). Aspect switching is explained by a switch in the representational memory traces which the currently seen object primes. Because similarity priming depends on how our low-level representation is structured, and because structuring might be modulated by voluntary attentional control, then aspect switching can be voluntarily controlled too. We *know how* to attend in specific ways so that the picture is seen as a  or as a  .

Crucially, aspect representation through schematization thus does not depend on attributing any specific concept to the perceived picture. I here join with Avner Baz's interpretation of Wittgenstein's aspects (here talking about seeing a particular face as looking like another):

> The perceived, dawning similarity of one face to another, I've proposed, is an internal relation between the two faces—a way of bringing individual things together perceptually without (yet, or necessarily) subsuming them both under some one concept, or generality that, as such, transcends them and any other finite set of its instances. What the two faces share is a perceived, experienced, physiognomy, not a determinate, empirical concept of which they may both be judged to be instances. (Baz, 2020, p. 116)

Such point generalizes beyond the mere fact of seeing adjacent faces as similar, to seeing schematization. When one is suddenly struck by the fact that a face is looking like Mary's, a new aspect is dawning on that perception: one now sees that face as 🙂 (imaging that 🙂 stand for Mary's face's aspect). The sudden representation of this aspect might be explained by the fact that one at some point comes to schematize in a way that produces a new similarity priming between the perceived face and perceptual memories of Mary's face. This explanation is not equivalent to an explanation where one would at some point apply the concept MARY'S FACE (or simply MARY) to her percept. Double dissociations are easily described. One could see a face as having a Mary's face's aspect (🙂) while refraining from judging that 'it is MARY'S FACE' (or 'it is MARY'), because one knows that this cannot be Mary. Parallelly, one could know that this is MARY'S FACE (or MARY) without seeing that face as 🙂, because Mary just had a massive facial plastic surgery operation, or simply because one sees Mary for the first time and has yet no memories of her face.

We already encountered a very similar case when I presented Goodman's challenge with **Figure 3**. We can see some particular objects as similar to different objects, depending on whether we attend to their color or shape. With the psychological apparatus described in the previous chapter, we can now have a precise understanding of what is happening. When we attend to the color of the objects, we structure our higher-level perception of such object so that color is representationally emphasized. This structuring in turn modifies similarity priming: one primes specific *color* memory traces (e.g., a simulated prototype color). Objects A and B are thus priming similar representations that are red, or a similar red-prototype. *They are seen under the same color aspect*. Both A and B are seen as red. They are thus perceptually grouped together following the classical gestalt principle of similarity. Likewise for C and D which are seen as having a similar blue aspect.

Now, if one attends to shape features, one will prime specific *shape* memory traces (e.g., a simulated prototype shape). Objects A and C are priming similar representations that are triangles, or similar triangle-prototypes. *They are seen under the same shape aspect*. Both A and C are seen as triangles. They are thus perceptually grouped together following the classical gestalt principle of similarity. Like for B and D which are seen as having a similar ellipse aspects.

The crucial point here is the following: low-level properties do not *on their own* determine schematization, and thus under what aspect an object is seen. *Of course*, the low-level color property of A (a determinate shade) is more similar to that of B than that of C or D. But the low-level shape property of A (a determinate shape) is more similar to that of C than that of B or D[64]. The point is that what A is similar to is underdetermined by its

[64] It is important to keep apart the low-level determinate color shade of an object ($\underline{\text{shade}_A}$) and the higher-level color *tout court* of that object (red). I take the latter to be an aspect, and to depend on schematization. Obviously, we often perceive an object *both* as having a low-level determinate shade *and* as having a color aspect. One thus perceives A as $\underline{\text{shade}_A}$, red. Strictly speaking then, I should accept that *colors tout court* (but not determinate shades) are high-level perceptual contents. I don't see this as an issue for my own defense of

low-level properties. Structuring solves such underdetermination. Structuring allows the perceptual system to prime specific representations by determining respects for similarity.

In this case, schematization is made explicit by putting similar objects adjacently: A and B are grouped together when they are seen as red, and A and C are grouped together when they are seen as triangles. Objects can come to be seen under different aspects. To repeat, the point is not so revolutionary in the case of **Figure 3**, since the grouping effect might be obtained merely by differentially attending to low-level properties of objects, like attending to their low-level determinate color shade (shade_A is more similar to shade_B than to shade_C or shade_D; shape_A is more similar to shape_C). But **Figure 3** serves to make phenomenologically striking the point that, through schematization, we can perceive different aspects of the same object. This is true even when objects are perceived in isolation: through schematization, one could see A as red, as triangle, as red, triangle, or as neither (if one does not schematize A). Through different schematizations, one sees A as being similar to different other things. Such similarity in most cases is not made explicit by similar adjacent stimuli but is an implicit similarity to perceptual memory traces.

This is the most important lesson I draw from Wittgenstein, and that I develop within a psychological framework (this move of course being a clear betrayal of Wittgenstein's anti-psychologism). One and the same object (bistable figures, but actually any possible object), can be perceived under different aspects, which amounts to perceive that object as (most of the time, implicitly) similar to different other objects. Such capacity is not akin to applying a concept to what is seen. Schematization relies only on the nonconceptual processes of structuring and similarity priming. Seeing aspects is not akin to thinking, a view I think remains faithful to the phenomenon Wittgenstein was chronicling.

My view of aspects as produced by schematization implies that aspect perception must be ubiquitous in perception, since schematization is part and parcels of normal higher-level perceptual processing. Aspect perception is thus not limited to very specific cases like bistable figures. We constantly and automatically see things as having some particular

---

Liberalism. I tend to see it as a strength: it keeps with the plausible intuition that we see color determinables (red, blue, green…) in addition to color determinates, while keeping with the also plausible intuitions that seeing color determinables is not dependent on possessing color concepts (RED, BLUE, GREEN…). I think an exactly parallel point applies to *shape*: one should distinguish between low-level shape determinates (maybe we should call them not shapes, but shape-gestalts) like shape_A (which is different from shape_C), from determinable shape aspects like triangle. Like for colors, we most often perceive objects both as having a determinate shape and a determinable shape aspect. We perceive A as shape_A, triangle. I also think that we should take shape aspects like triangle to be high-level perceptual contents as opposed to their determinate counterparts. But I am *not* saying that this stands as a good argument for Liberalism. A Conservative could claim that if I want to name them "high-level content", so be it, but colors and shapes were never at stake in the debate anyway. Fair enough (that's why I focus on defending *kind* aspects in what's to follow), but I take it as a great strength of my account that we will end up with a much more precise distinction between low-level and high-level contents, even though it doesn't draw the border in the same way as traditionally drawn in the admissible contents of perception literature.

aspects, that is, as being implicitly similar to things we've seen before. This idea that aspect perception is constitutive of normal perception is interestingly attributed to Richard Wollheim by Avner Baz:

> Wollheim and many others have felt that the dawning of Wittgensteinian aspects is, *must be*, revelatory of (normal) human perception *as such*—of what Wollheim calls 'straightforward perception'. Specifically, these philosophers have come to hold one version or another of the idea that, over and above the lighting up of aspects, there must also be a continuous version to the perception of aspects, and that all (normal) human perception can, and ought to, be understood as the perception of aspects. (Baz, 2020, p. 94)

Interestingly, Wollheim thought that aspect perception necessarily depended on concept application, leading him to conclude that perception was ubiquitously and constitutively conceptual in nature. I obviously reject this view. On my view, seeing aspects is ubiquitous, but still a *nonconceptual* representational capacity carried out by schematization.

There might be some exceptions to the ubiquity of aspect perception. Maybe what painters do when they reproduce a landscape is to try to focus on the field of low-level determinate properties that they perceive, and to detach themselves from seeing anything under any particular aspect. They want to reproduce *that particular landscape*, so they should really try to detach themselves from being sensitive to similarity associations to what they have seen before. This is a somewhat metaphorical description though, and one might doubt whether stopping oneself from seeing aspects is even psychologically possible, considering their automatic nature. As I said in note 64 (above), I think *colors* should be taken as aspects of determinate shades. I think this makes aspectless perception almost impossible, since even a uniform color field will be seen as having a color aspect.

A potential candidate for aspectless perception might be the perception of very young infants (who have not much perceptual memories stored). But it is plausible that some schematizations are evolutionarily hardwired (cf. VI.1 AUTOMATION). Aspectless perception might also happen for the simplest perceiving organisms with no schematization capacities.

I leave these questions open in this work. Even if aspectless perception is possible, it remains the case that it must be the exception rather than the rule, especially for adult human perceivers.

V.1.2. A HIERARCHY OF ASPECTS

Another point that I wish to emphasize is that the same object might be viewed under various aspects *at the same time*. There exists a *hierarchy* of aspects in perception. For instance, one might see something as having not only a duck aspect, but also an animal aspect (a "superordinate" aspect), and, if say one is a trained ornithologist, as having a *Mergellus albellus* duck aspect (a "subordinate" aspect). This possibility is explained by the possibility of differentiated hierarchically organized schematization processes in perception.

One can think of hierarchical schematization by imagining yet again some perceptual space with clustered manifolds. Hierarchical schematization happens because manifolds might be embedded into higher-level manifolds. The duck representational manifold might thus be embedded into the higher-level animal manifold. There might thus be similarity primings at different ordinate levels for the same perceived particular. One might prime representations that are in some ordinate, "basic-level" manifold, or some other embedding (superordinate) or embedded (subordinate) manifolds.

One could also choose to go against simultaneity views and maintain that we can only see one aspect at a time (but of course, one could still see different aspects at *different* times). This view seems less intuitive. Seeing something as having a Fiat 500 aspect does not seem to impede one from seeing its car aspect. One could even claim that one just a priori *cannot* see a Fiat 500 aspect without thereby seeing a car aspect.

But we should beware of such intuitions. Because we can perceive different aspects for one and the same particular, it might seem that we can perceive all these aspects simultaneously, even though we actually only alternate between perceiving these aspects. Thus, on this view, many perceptual experiences would be multistable like for the duck-rabbit image, though it would be a case of *vertical* multistability between compatible aspects, not a case of *horizontal* multistability between incompatible aspects as in the duck-rabbit case. It's just that in everyday cases of multistability, since the aspects between which we switch are compatible, these aspect switches are not easily noticed. We seamlessly and continuously (and maybe very quickly) switch between seeing something as having a car and a Fiat 500 aspect. What's abnormal about *ambiguous* figures is precisely that the different aspects between which we switch are *incompatible* (nothing can be both a duck and a rabbit), explaining why we notice them.

As interesting as they are, these points about perceiving different aspects, compatibly vertically organized or horizontally incompatible, simultaneously or serially, should not have any impact on considerations about the admissible contents of perception. In fact, suffices that *one* aspect of an object be perceived at a time to count as a kind of high-level perceptual representation, and to vindicate Liberalism. It is high time to say more about the high-level nature of aspects.

V.2. Aspects: First Form of High-Level Representations in the Mind

V.2.1. The Representationally Unreducible Nature of Aspects

I think aspect perception is a genuine high-level representation in the sense that it does not reduce to structuring its low-level feature representations in a specific way. Aspect perception *does* depend on structuring, but it also depends necessarily on similarity priming. Structuring-without-priming does not yield an aspect representation, but only a

new low-level (or maybe mid-level) representation, which *is* reducible to lower-level features. A representation of a determinate square-gestalt is reducible to representations of local edges, though perceiving a square aspect is not. Aspects introduce a new representational kind in perception which remains vertically articulated to lower-level perceptual representations (and ultimately, causally articulated to particulars in the world), but is unreducible to them (Skrzypulec, 2018). They classify particulars into equivalence classes of similar representations. They are a primitive type of *kind* representation.

Schematization would be impossible without some representation to schematize in the first place. Structuring and similarity priming require that one represents the low-level properties of an object. It is this structured low-level representation that primes similar representations from perceptual memory. Aspect representations do not erase the low-level representation from which they are stemming. Take the case of seeing an object *o* as having an apple aspect . This aspect representation requires that one has already represented an ensemble of low-level features for *o*: *o* is seen as shape_{42}, color_{85}, texture_{22}, size_{58}… (at specific spatial positions). Many of these low-level representations are discarded, deemphasized, or bounded in more complex features through structuring (so that they do not semantically confuse similarity primings). But the lower-level, unstructured feature representations themselves are still tokened. When one perceives the apple aspect of *o*, one deemphasizes the texture and the color representations of *o* through structuring. But still, and quite obviously, one does not stop perceiving some determinate color and determinate texture of *o*. One sees both low-level properties of *o* and at the same time sees o *as being similar to other things with respect only to shape*.

The aspect representation remains necessarily tied up to the low-level representation: it's always with respect to some low-level feature representations that one perceives an object as being similar to some others. Thus, aspects are vertically articulated to low-level representations. Notice that such articulation might be indirect. To perceive a cat aspect, one might first be required to perceive some cat-shape aspect and some cat-texture aspect. Remember from note 64 that I take classificatory determinable shapes and textures to be themselves aspects, as distinguished from determinate shapes and textures. One must distinguish between the determinate shape_i feature of a particular cat, from the cat-shape aspect (which is shared by cats, cat silhouettes, cat sculptures…), and the cat aspect (which might depend on tokening cat-shape and cat-texture aspects). The latter aspect representations depend on schematization processes, while the former determinate representation does not. In any case, the fact that representing a cat aspect depends itself on representing some cat-shape, cat-texture and cat-size aspects does not threaten the idea that the cat aspect indirectly depends on representing some low-level features. Aspects are necessarily *vertically articulated* to low-level determinate representations.

Still, aspects remain *unreducible* to the low-level representation to which they are articulated to. This should again be obvious if we notice that representing *o* as shape_{42}, texture_{95}, size_{52} is different from representing *o* as shape_{42}, texture_{95}, size_{52}; similar to shape_{41}, texture_{96}, size_{51}. One could well see a cat but fail to see it as similar to other things (i.e., fail

to schematize it under a cat aspect). This could be due either to the fact that one doesn't structure her low-level representation in a proper way, or because one lacks similar memory traces. One could for instance fail to see that a face looks similar to that of Donald Trump, even if one has many memory traces of Donald Trump. This might be because one is paying attention to the wrong features (e.g., nose shape), thus failing to prime relevant memory traces. The only memory traces that are primed are undifferentiated memories of "male faces". One thus sees a male aspect but not a Donald Trump one. Then, someone might point to the relevant feature to attend to (e.g., eye distance and hair color), thus producing a structured representation that does prime differentiated memory traces corresponding to Donald Trump. When this occurs, an aspect switch à la Wittgenstein is experienced: one switches from seeing the face as having an undifferentiated male aspect, to seeing the face as having a Donald Trump aspect (or one *adds* a Donald Trump aspect representation on top the male aspect representation, if it is possible to perceive various aspects simultaneously). Crucially, after this aspect switch, the low-level representation might have remained the same. Thus, while aspects are vertically articulated to and supervene on low-level representations, they are not reducible to them.

A further reason to believe in the non-reducibility of aspect representations to low-level representations stems from the fact that two different individuals could perceive the exact same low-level features but still not the same aspect. Someone who has never seen Donald Trump before will obviously be unable to perceive a differentiated Donald Trump aspect when seeing Donald Trump or a person who looks like him, because one does not store any representation of Donal Trump-looking faces. Thus, we can conceive of two subjects having the exact same low-level visual representation (including determinate gestaltic mid-level representations), but different aspect representations. One sees merely a male aspect, and the other sees a Donald Trump aspect (or maybe both at the same time). One will thus recognize what she sees as a male face, while someone else will recognize it as the face of Donald Trump. Their perception come under different equivalence classes.

Notice that this is true *even if both individuals pay attention to the same dimensions of the face* (e.g., to the hair shape and color) *and have similar bottom-up untangling processes*: their schematization will still be different because these evenly structured representations will not prime the same type of perceptual traces. The conclusion is that aspect representations are not reducible to low-level (and "mid-level") representations. They are good candidates to be genuinely high-level representational classifications in perception.

V.2.2. No Luck in Aspect Perception

The fact that aspect representations can diverge even keeping constant top-down attentional selection and bottom-up untangling is crucial in avoiding undesirable cases of *untangling* or *attentional luck* in aspect perception. One could well attend, by sheer luck, to the same features of Donald Trump that allows *me* to schematize a specific Donal Trump aspect, *without thereby schematizing a Donald Trump aspect*. This idea is even more vivid in

comparisons of perceptual experts with novices. A novice radiologist might just structure an x-ray image in the same way as an expert radiologist, but only the expert radiologist would perceive tumor-aspects. In fact, only the expert radiologist has structured memory traces of tumors on x-ray images, so only such radiologist can schematize in a way that perceptually reveals tumor aspects (Stokes, 2021a). We do not want to conclude that the novice luckily sees a tumor aspect simply because she attends by sheer luck to the image in the same way as the expert radiologist.

This point is important, because it puts to doubt a recent defense of perceptual Liberalism proposed by Madeleine Ransom, who claimed that *expert patterns of attention themselves* constitute kinds of high-level representations (Ransom, 2020b, 2020a). For Ransom, experts learn to see new structural relations between low-level properties. They can learn to see new "gestalts". For instance, an ornithologist can become visually sensitive to a specific gestaltic organization typical of *wrens*. Furthermore, if her goal is to identify wrens, the ornithologist can attentionally weigh this gestalt so that it will be salient in her visual representation. For Ransom, such structured and goal-directed attentional weighting is sufficient to claim that the ornithologist *sees* the high-level property *being a wren*. As she puts it:

> The systematic encoding of second-order relations between these low-level properties, as well as the evidential weighting of such properties, is sufficient to diagnose in perception membership in a high-level category for most exemplars of that category. (Ransom, 2020a, p. 245)

For Ransom, specific "patterns" of attentionally weighted low-level features are sufficient for high-level visual representation: "[i]nstances of high-level properties are instantiations of a given pattern" (*ibid*., p. 244). This is so because attentional weighting patterns are precisely produced to allow for specific categorizations:

> The information contained in the learned attentional weights is not simply 'look at this property' but is rather best understood as something akin to 'this property is important to x degree in confirming or rejecting this exemplar as a member of a given category y'. (*ibid*., p. 241)

I think Ransom's view is facing a dilemma.

On the one hand, we can take it that attentional patterns represent some kind of high-level property because they are deployed to be *consumed* by specific processes of categorization. For example, a specific attentional pattern *P* might be intentionally and expertly deployed over an object *o* because *P* selects properties that are evidential for *wrens*. In that case, one might be tempted to say that *P* grounds representations of *wrens*. When one deploys *P* over *o* to discriminate wrens, and when *o* does possess the relevant wren-diagnostic properties, one thus sees *o* as a wren. This interpretation would allow to avoid issues of attentional luck because the novice who attends to *o* like the expert by sheer luck would *not* be deploying attentional pattern *P* as a way to recognize wrens. The novice would thus not see *o* as a wren.

The problem with this view is that it presupposes a questionable psychosemantics, one that would have as a rule that a representation *R* can be said to represent *x* if (1) *R* is evidential of *x* and (2) *R* is intentionally deployed so as to categorize *x*. The main problem with this view is that it presupposes the questionable premise that a representation's content is determined by how that representation is used. In other words, it presupposes a consumer-oriented teleosemantics of perceptual representations. As I have developed in II.3 PSYCHOSEMANTICS, such kind of argumentative strategy for perceptual Liberalism is unsatisfactory, because it relies on (psychosemantics) premises that might be just as polemical (if not more) as Liberalism itself. Furthermore, it seems this consumer-oriented position would just end up trivializing perceptual Liberalism. After all, it is undeniable that perceptual processes are inputs to some categorization processes. Even perceptual Conservatives accept this (since they surely accept perceptually-based categorizations). But if we were to attribute contents to perceptual states based on the kind of evidence they give for some categorization decision, then even a Conservative would need to accept that perception trivially represents high-level contents. As I have argued in I.2 SEEING AS SEEING-AS, the admissible contents of perception debate is made philosophically consequential only if we strictly separate perceptual content proper from perceptually-based content (most importantly: perceptually-based categorization). It seems Ransom's view threatens this separation.

A second interpretation would be to view patterns of attentionally weighted properties as themselves being high-level representations *whether or not they are deployed to serve specific categorization.* The advantage of this view is that we do not need to appeal to further controversial psychosemantical claims. The obvious issue though is that this view has trouble accounting for cases of structuring luck: a novice attentionally weighing and structuring his perceptual representation of a bird in the same way as an expert ornithologist, but by sheer luck, would have to be said to perceptually represent a *wren*. This result is troubling because the attribution seems totally arbitrary. After all, the novice has no idea what wrens are (he has never heard of them before), and most importantly, the novice associates no specific behavioral generalization to such structuring pattern (it is applied luckily and randomly). We should warry of not attributing high-level representations (wren for the novice) which have no bearing on behavior.

My own defense of Liberalism, relying on schematization, can integrate Ransom's insistence on attentional selection (and perceptual structuring in general), but adds the fundamental ingredient of *similarity priming*. Without it, attentionally-modulated structuring is helpless in constituting some kind of high-level content. It is the combination of the two that grounds high-level content in the form of *aspect* representations. Similarity priming needs perceptual structuring (guided by attentional selection) so that the representations that are primed are behaviorally and semantically relevant (and not just similar representations on irrelevant respects). Structuring needs similarity priming to yield genuinely high-level, nonreducible representations in perception.

Some critiques have argued that bistability, phenomenal contrasts and phenomena of perceptual expertise in general (e.g., learning to recognize pine trees or tumors on x-ray images) can be accounted for solely by appealing to attentional effects (Nanay, 2011; Prinz, 2013). Since such attentional effects are on their own merely modifications of low-level representations, we cannot interpret bistability and expertise phenomena as vindicating Liberalism. While Conservatives are right to point that attentional or gestalt effects are insufficient to ground high-level content (*pace* Ransom), they are wrong in being confident that this is enough to reject Liberalism in favor of Conservatism. They overlook the idea that attentional effects can be bound to a further psychological mechanism: similarity priming.

This additional step is easy to miss because it certainly intermingles with the first at an introspective, phenomenological level. When we switch from grouping objects by color or shape in **Figure 3**, or switch from the duck to the rabbit, we both change attentional weighting pattern *and* similarity priming. In that case, it is precisely because one changes her attentional weighting pattern (through top-down control) that one changes the similarity priming process, and thus the particular aspect that is being seen, red-aspect or triangle-aspect; rabbit-aspect or duck-aspect.

The difference between these two steps can be intuited by considering cases where only one of them occurs. One can change her attentional focus without changing the aspect that is being seen. Imagine someone who has never seen and knows nothing about ducks. Upon seeing the duck-rabbit image, she thus only sees a rabbit aspect. In that case, it is at least conceivable that by sheer attentional luck, such a person could be attending to the picture in the exact same way as me when I am seeing the duck aspect. In this scenario, there would be a phenomenal switch between the two moments (since different patterns of attention plausibly correspond to different perceptual experiences), but *not* a switch in aspect perception (all along, the person is perceiving a rabbit aspect, though through different attentional patterns).

Oppositely, one can imagine a situation in which attentional weighting patterns remain fixed, but aspect perception changes. This happens if we compare the perceptual experience of one who just attends luckily to a golden chanterelle like an expert mycologist, without having ever seen one, with an expert mycologist who attends in the same way but for whom this particularly attended perception primes specific mnemonic traces. This mycologist might be said to have the exact same low-level structured perception, but still her whole perception will be different since she has a further (high-level) perceptual effect in the form of similarity priming. In other words, these two people do not *schematize* in the same way what they are seeing, though they *structure* what they see in the same way. They see different aspects (maybe the lucky novice sees only a general mushroom aspect, while the expert sees a golden chanterelle aspect). Here again, we get a perceptual difference between the two cases, though this perceptual difference is not the result of different patterns of attention. I thus concur with Susanna Siegel in her view that, *even keeping attentional patterns fixed*, a novice might only see some low-level property while an

expert might additionally see some high-level property (Siegel, 2006, 2011). Although, as we will soon see, *contra* Siegel, I do not think that the expert literally sees a *natural kind* property that the novice doesn't.

I think the reason why phenomenal contrast arguments are so controversial, with Liberals and Conservatives proposing equally plausible interpretations (see II.2 PHENOMENAL CONTRASTS), is because *both* camps get something right. Conservatives are right to point that, when we learn to recognize a new class (e.g., recognize *pine trees*), there is usually a change in *low-level* phenomenology due to changes of attentional patterns (the novice and the expert are expected to attend differentially to the tree, excepting cases of attentional luck). But Liberals are also right to point that the phenomenal contrast is *not* limited to a change in low-level, attention-based phenomenology. There is also a phenomenal change that is to be further explained by a change in high-level content: the novice sees only a tree aspect, while the expert additionally sees a pine tree aspect. The expert sees something as looking like other *pine trees* he's seen before (remember this representation is neither conceptual nor explicit or conscious), while the novice sees that same thing as looking like undifferentiated *trees* he's seen before.

Phenomenal contrast arguments as they stand do not easily allow to adjudicate between Liberalism and Conservatism, because they describe a contrast which mingles both a low-level attentional contrast *and* a high-level aspect contrast (the same goes for bistable figures). My view thus endorses Liberals in their basic assumption that there is a phenomenal contrast between the novice and the expert, but I insist that it remains difficult to determine what drives the effect between low-level or high-level representations. That's why my own non-phenomenological, empirically-based, bottom-up psychological approach is I think superior. It just makes psychologically plausible the idea that the difference between the novice and the expert is not just one in attentional skill and modifications of low-level representations, but also one in *mnemonic* background, similarity priming, and thus ultimately *aspect* representation.

V.3. INDIVIDUATING ASPECTS AND THEIR ACCURACY CONDITIONS

V.3.1. ASPECTS OF THE WORLD, NOT ASPECTS OF ONE'S MIND

Schematization vindicates the idea that we can perceptually represent some specific high-level contents: *aspects*. Aspects are grounded in the associated tokenings of a perceptual representation and perceptual memories. This association is guided by a principle of similarity: the perceptual memories that are primed are those that are maximally similar to the occurrent structured perceptual representation. Structuring ensures that such similarity priming is possible, by determining *respects* for similarity, and further ensures that the memory representations that are primed are semantically and behaviorally coherent. This

process of schematization produces high-level representations because the web of tokened similar representations that constitutes an aspect is not reducible to the occurrent low-level representation from which it derives.

I need to say more about *what* is represented through aspect representations[65]. What exactly is aspect content? And what properties are perceptually referred to or denoted through aspect representation? What kind of property is attributed by seeing some object as having a certain aspect (or certain aspects)? What property would make an aspect representation *accurate* or *inaccurate?*

A first option I put aside is that aspects would represent purely *mental* properties, or that aspects would be purely "meta-representational" or "meta-perceptual". One could be tempted to argue that aspects represent that some representation is similar to some others. This similarity is purely *internal* to the mind: it is a mental property. Through schematization, one would represent something about one's perceptual representations. Since schematization is a perceptual process, one would thus *perceive something about one's perception*. Aspect representations would be something akin to "*seeing what I see* as similar to something I have seen before".

There are several reasons one should reject such meta-perceptual view of schematization. First, it is intuitively not convincing to presuppose that perception could be about itself. One surely can think things about what one is seeing. But it is odd to say that one *sees* things about what one is seeing. Moreover, schematization never invokes meta-representational capacities. It only invokes *association* capacities between an occurrent representation and mnemonic traces through similarity priming. These associated representations form an independent high-level representation, which is *not* a meta-representation. A drawing of a face is not a meta-representation of its parts (the eyes, the nose, the mouth). A drawing of a face associates different parts (which themselves are representations of parts of a face) to represent a new object (a face). Likewise, schematization associates[66] different representations of an object to represent a new representational property of this object. This new representation does not itself represent a specific relation between representations. It simply represents a new property of the object. If it did represent a relational, meta-representational property, we could reasonably doubt that schematization would be a genuinely *perceptual* capacity.

I don't think schematization is a form of metacognition either, at least not as it is usually understood in the psychological literature: it is not a form of monitoring and control of one's own mental processes. Schematization is just a particular high-level perceptual process, just like lower-level processes such as constancy computations or feature binding. Of course, schematization *can* be monitored and controlled through metacognition. One

[65] Obviously, asking this kind of question takes us far away from the anti-representational positions of the Wittgenstein of the *Philosophical Investigations*.

[66] This association is of course a *mental* association, that constituted by similarity priming. It is not a *spatial* association like parts in a drawing or a *temporal* association as in a musical melody.

might be more or less confident that what one is seeing really has some aspect or not (you might see something as having a rabbit aspect but feel unsure about it, maybe because you see it in a very dark environment). Also, you might have some metacognitive feelings caused by perceptual schematization, even though you cannot explicitly judge what you have schematized. For instance, you might metacognitively feel that what you are seeing looks like something you've seen before because you implicitly prime similar representations of it from memory, though you cannot think about what it is that it looks like. You thus have something we could call a recognitional tip-of-the-tongue experience. You *feel* you've seen something like it before, but you cannot *think* about what it is[67]. Also, metacognition is usually thought to depend on executive prefrontal regions (Fleming & Dolan, 2012), while schematization isn't (it only depends on structuring and priming capacities usually thought to be realized in the ventral temporal stream). So, while schematization can be monitored by metacognitive capacities and can yield metacognitive feelings, it is not itself akin to a metacognitive process. Schematization represents properties of the world, not properties of one's mind.

V.3.2. Aspects Cannot be Natural and Functional Kind Representations

Then, we can ask: what kind of worldly property does the association of occurrent perceptual representation with perceptual mnemonic traces represent?

One possibility is that aspects represent *natural kind* properties, such as the *applehood* of apples, the *cathood* of cats, or the *pinetreehood* of pine trees. Schematization might denote natural kind properties because one can reasonably suppose that similarity relation between objects is grounded in their sharing of a natural kind property. For instance, the fact that all apples are similar (in certain respects) is explainable by the fact that they all share a similar genetic code (and similarly for cats and pine trees). Apple aspects would thus denote applehood, the natural kind property that determines the similarity relation in the current schematization. This idea can extend beyond natural kind properties to *functional* kind properties. What makes all *chairs* similar to each other (in certain respects) is the fact that they fulfill the same function (e.g., being sit-able by adult humans). Chair aspects would thus refer to *chairhood* (understood functionally). Aspects would be on par with concepts in that they could refer to something like deep "natural essences" or "functions" in the world[68].

The idea that aspects refer to such deep natural or functional kind properties is in line with much current defenses of perceptual Liberalism in the philosophical literature (Bayne, 2009; Siegel, 2006, 2011). Schematization would provide a psychological

[67] "Déjà-vu" feelings might be thought of as the illusory counterparts of such recognitional metacognitive feeling (maybe because the schematization process has gone wrong).

[68] That aspects represent such hidden properties does not mean that we explicitly know what they represent. Just like for concepts, the reference of an aspect might be opaque to the subject. One might not know that apple aspects refer to the *applehood* biological property, just like one might ignore that the concept WATER refers to H_2O (on Earth).

mechanism that grounds such representational capacity. Nonetheless, I do not think this is the right view for the contents of aspect representations. The main intuition against aspects representing natural or functional kind properties is the following: visually representing an apple aspect when seeing a *plastic apple* should count as an *accurate* aspect representation. It is accurate to represent the plastic apple as instantiating an apple aspect, while it is false that the plastic apple does not instantiate the *applehood* biological kind property.

When one is presented with a plastic apple, there is a sense in which one is warranted in saying that "this is an apple", and a sense in which one isn't. I think the sense in which one is warranted and accurate in saying that "this is an apple" is the *aspect* sense (it could be translated as "this has an apple aspect"). The sense in which one isn't warranted and accurate in saying that "this is an apple" is the *natural biological kind* sense.

It is rational for a monkey to grasp a plastic apple and try to eat it. The monkey's 🍎 aspect representation is *accurate*, and the behavioral repertoire associated with it is thus rationally potentiated. On the other hand, a human adult who would think of a plastic apple that 'this APPLE would make for a good APPLE PIE' would be wrong, since the general concept APPLE refers to the apple fruit, not to plastic apples[69]. Since upon seeing a plastic apple, one is accurate in tokening 🍎 but inaccurate in tokening APPLE, we can conclude that 🍎 does not have the same content as APPLE (though see note 69 for a crucial caveat). 🍎 does not represent the natural kind property *applehood*.

Another way to drive the same point home is to think about a twin earth scenario for aspect representations. If aspects represented natural kind properties (like the natural kind concept WATER), then we would expect that an apple-aspect representation on Earth (🍎) would represent something different than its twin aspect representation (twin-🍎) on Twin-Earth. Like in the classical Putnam scenario (Putnam, 1975), let us imagine two protagonists, Mary, living on Earth, and Twin-Mary, living on Twin-Earth. Earth and Twin-Earth are exactly alike except for the chemical composition of one of their fruits: apples on Earth have a genetic code ABC while twin-apples on Twin-Earth have a genetic code XYZ. Notice that this genetic difference does not produce visible differences between

[69] An important caveat has to be made here. Since I adopt a "semantic pointer view" of concepts (see III.1.1 THE NON-CONCEPTUALITY CRITERION), I think that the concept APPLE can point to differentiated kinds of information depending on the context of tokening, and thus can have context-dependent reference. Thus, one might think that 'APPLES are BEAUTIFUL', where the concept APPLE points only to perceptual information (in that case, the concept APPLE can be said to be a purely observational concept), and thus refers to fruit apples and plastic apples altogether. In another context, one might think that 'APPLES are HEALTHY', where in that case the concept APPLE points to some non-perceptual information, such as the nutrition values of apples. In that case, the concept APPLE only refers to fruit apples. Ultimately then, I think there is no strict dichotomy between observational and non-observational concepts, but only a difference in context of tokening. The same concept might sometimes be used observationally, and sometimes non-observationally, changing its reference and thus semantics (Quilty-Dunn, 2021). In most of what is to follow, I will understand concepts as being used in *non-observational* ways, i.e., as also pointing to non-perceptual information, which allows them to refer to natural or functional kind properties. This strongly distinguishes concepts from aspects.

apples and twin-apples. They are exactly alike in all superficial respects (they are roundish, mid-size objects which are edible, taste sweet and slightly sour etc.). Mary and Twin-Mary are exactly alike, both physically and in terms of their experiential history. Most importantly, they have had the exact same perceptual experiences of apples/twin-apples.

In this scenario, Putnam's point still holds: Mary's natural kind concept of APPLE does not represent the same property as Twin-Mary's natural kind concept APPLE. The former refers to fruits with the ABC genetic code (apples), while the latter refers to fruits with the XYZ genetic code (twin-apples). When Mary goes exploring the universe, lands on Twin-Earth, and says "apple!" (expressing APPLE) while seeing a twin-apple, *she is wrong* (since her own Earthly concept of APPLE refers to ABC-fruits, not XYZ-fruits).

What about Mary's 🍎? Is Mary wrong when, landing on Twin-Earth, she tokens the perceptual representation 🍎, i.e., tokens a representation of an object as being perceptually similar to other things she has seen before in certain respects? I think the answer is negative. Mary is *right* in tokening 🍎, even on Twin-Earth. After all, the twin-apple *does* look like things she has seen before under certain respects, namely, Earthly apples.

Notice that the schematization processes going on in Mary's and Twin-Mary's perceptual systems are exactly type-alike when they see the same twin-apple in the same viewing conditions: their low-level visual representation is the same, they structure in the same way, and they prime exactly the same perceptual mnemonic traces. This result is given by the assumption of the thought experiment that Mary and Twin-Mary are exactly alike psychologically and physically, and that they have the same perceptual history (i.e., the same perceptual memory traces). The schematization process is only guided by similarity along perceivable feature dimensions, *not* by similarity along non-perceivable features (e.g., chemical properties). The fact that Mary and Twin-Mary are concurrently right in tokening 🍎 upon seeing the twin-apple shows that this aspect representation does not refer to the natural kind properties of *applehood* or *twin-applehood*. If it did, we would need to conclude that Mary is wrong, and that Twin-Mary is right in tokening 🍎. But this does not come as an intuitive conclusion, since we want to say that *both are right*. Both apples and twin-apples are correctly represented by the "apple"-aspect.

Why are aspects different from concepts regarding their representational content? Why do Mary's 🍎 and Twin-Mary's 🍎 represent the same thing (have the same accuracy condition), while Mary's APPLE and Twin-Mary's APPLE concepts do not? The answer lies in the difference in representational capacities between perception (to which aspects pertain) and cognition (to which concepts pertain). Concepts in cognition can[70] be decoupled from the demonstrative perceptual relation of the subjects to particulars (Burge, 2022). One can think about apples with one's eyes closed or without tasting one. What non-observational concepts refer to can thus be opaque to the subject. One might not

[70] The modal precision is crucial here since cognition can also include demonstrative thoughts.

realize that her concept of APPLE refers to a particular fruit with a particular chemical composition. Non-observational concepts point to a lot of non-perceptual, non-demonstrative information that allow them to represent non-perceivable properties of the world. Putnam-like Twin-Earth scenarios work because the natural kind concepts of APPLE on Earth or Twin-Earth are taken as representations which refer to some invisible micro-property of some object[71].

Also, concepts are *public* symbols: their reference is determined by the idiosyncratic practices of a specific linguistic community. As Burge has famously pointed out (Burge, 1979), one would be *wrong* in thinking that she could have ARTHRITIS in the thigh, even though one (falsely) believes that ARTHRITIS is a pain in the thigh. This is because determining what ARTHRITIS refers to is not up to the individual tokening the concept. What ARTHRITIS represents is determined by the linguistic community of English users (and especially by the authoritative subgroup of medical practitioners). Likewise, what APPLE represents is determined by the same community of English users. My own concept of APPLE could point to representations as of being an edible orange and sour fruit, I'd still be wrong in attributing APPLE to an orange.

On the other hand, aspects cannot decouple from their *representatum* like concepts do. Aspect representations (like all perceptual representations) constitutively depend on the demonstrative perceptual relation between the subject and the object. One cannot represent 🍎 without having ones' eyes opened and receiving sensory input from some object. One needs to perceptually represent some low-level properties to get schematization started. Because of this, aspect representations, contrary to conceptual concept representations, are referentially *transparent*. Transparent not in the sense that we are explicitly conscious of what aspects refer to (if we were, we wouldn't have an admissible content of perception debate in the first place). But aspect representations are referentially transparent in the sense that they refer to a property we are currently *perceiving*. Aspects cannot refer to non-perceivable properties. They cannot refer to the chemical composition of an object, which we cannot perceive. Furthermore, the accuracy condition of aspect representations is independent from social linguistic normativity. Aspects are not language-of-thought-like, public symbols which are given a referent by social practice. Aspects are perceptual representations for which reference is provided by occurrent perceptual relation to the perceivable environment and internal perceptual capacities. Aspectual content, unlike conceptual content, is thus impervious to twin-earthing *à la Putnam* or to social determination *à la Burge*. We should thus not take aspects to represent *invisible* properties such as natural kind or functional kind properties. But what do aspects represent then?

[71] Again, this is only true in a context where APPLE is used non-observationally as a natural kind concept, i.e., a semantic pointer pointing to microphysical (chemical) information. The *observational* concepts of APPLE on Earth and Twin-Earth on the contrary refer to the same (perceivable) properties.

V.3.3. Objective Similarity and Superficial Kind Properties

The last points shouldn't bring us back to thinking that aspects represent a relation between representations. Aspects refer to the world when they attribute a property. If this property is not instantiated while the aspect representation is tokened, then the aspect misrepresents. What is this property attributed by aspects, which is neither a low-level property, nor a natural (or functional) kind property, nor a metarepresentational property?

The answer I recommend is the following: aspect representations attribute an *objective, worldly relation of similarity* between the currently perceived object and a particular group of objects. In other words, aspects represent an object as pertaining to a group of similar objects. What I represent when I represent an object *o* as 🍎 is the *fact* that *o* is similar under certain respects with a particular group of other objects. This group of objects is the one represented by primed perceptual memories. I talk of an "*objective* worldly relation of similarity" to insist that we are not talking of some mental, metarepresentational similarity. We are talking of a similarity property between physical objects. A certain physical object *o* is represented as being similar to other objects o_1, o_2, o_3 (or to one prototype of these objects). There is an objective fact of the matter about similarity relations in the world. Features of the world, just like features in the mental space, can be organized in a multidimensional space, in which distance between objects determines their similarity. A $1m^3$ cube is objectively more similar to a $2m^3$ cube than to a $100m^3$ in the size respect[72].

One can thus be wrong about the world when representing aspects. One can misrepresent. Imagine that I perceive a certain object *o* as 🍎. Through this aspect representation, I represent *o* as similar to some other object o_1 (which turns out to be another apple, though remember that I do not here represent *o* as being an APPLE). Now, imagine that I also have stored in perceptual memory a trace of another object o_2, which is objectively more similar to *o* than o_1. For example, *o* might actually be an orange, which I (mis)represent as having an apple aspect (because maybe I am seeing it in a dim kitchen cupboard, and thus wrongly see it as having an apple-shape). To be accurate, I should actually represent *o* as having an orange aspect. Notice something crucial here: if aspects were only attributing metarepresentational properties, then I would be *accurate* in representing my current (inaccurate) perception as of an apple-shaped object as being similar to perceptual mnemonic traces of apples. I would be *accurate* in representing *o* (an orange seen as having an apple-shape) as an apple-aspect. But aspects are not attributing metarepresentational properties. They are attributing an objective similarity relation

[72] I include *relational* properties in the furniture of the objective properties of the world. This point is crucial because one might think that many perceived properties (especially "secondary qualities" like colors) are only relational properties (Shoemaker, 2003; Thompson, 2007b). I think we can keep a notion of *objective similarity* even though we are only dealing with similarity between relational properties. If *green* is the relational property of being normally disposed to cause green experiences in some perceiver, then two objects remain objectively similar in that they are both normally disposed to cause green experiences in some perceiver. Of course, this view would make properties represented by aspects themselves relational properties.

between objects. Through my apple aspect representation, I represent *o* as more similar to o_1 (an apple exemplar, or an apple prototype) than to o_2 (an orange exemplar, or an orange prototype), and this is inaccurate: *o* is objectively more similar to o_2 than to o_1, however I am representing *o*.

One remark is crucial here. If aspects represent objective worldly similarity, then aspects' accuracy conditions must come in degrees: one can be *more or less* accurate in attributing an aspect to an object. That's why it is better to talk of *accuracy* conditions instead of *truth* conditions for aspects (Crane, 2009). It is more accurate to attribute an orange aspect to *o* than an apple aspect, but it is more accurate to attribute an apple aspect to *o* than a chair aspect, or an Eiffel tower aspect. Accuracy conditions for aspect representations is thus not absolute. No aspect attribution is fully inaccurate, or fully accurate, because there is no similarity level that is nihil, or that is total (except maybe in cases of numerical identity. But aspect representations go beyond representing one and the same object as being similar to itself).

One can only aim at being maximally accurate relative to some standard. The standard can vary. One can be maximally accurate compared to one's own idiosyncratic capacities. If the only fruit Mary has ever seen in her life are oranges, then seeing an apple as having an orange aspect might be the maximally accurate aspect representation *for her*. The standard might also be taken at the social level. In the current social context, attributing an apple aspect to *o* which is an orange is representationally underperforming. Most people would more accurately be able to classify *o* as having an apple aspect. We could thus say that one is (relatively) inaccurate in attributing an apple aspect to the orange relative to an average social standard. Perceptual *experts over*perform since they can attribute even more fine-grained aspects (with higher similarity levels) relative to the average social standard. For example, perceptual experts would be able to attribute to *o* a subordinate aspect like a Fuji apple aspect.

Similarity between objects is an objective property of the world. It is such similarity which is represented by aspects. We are here talking about the only kind of similarity that the perceptual system is sensitive to, which we might call *superficial* similarity. Schematization associates a current low-level perceptual representation with perceptual memory traces through similarity priming. Low-level perceptual representations (including perceptual memories) only represent physical properties of objects that are visible, such as color, shape, or texture. These are properties which pertain to the surface bodies of objects. Oppositely, microscopic chemical properties or functional properties are not visible, they are not represented as low-level properties by the perceptual systems. They do not enter into schematization processes.

Just like objects with some microscopic properties (cats) instantiate some natural kind property (cathood) or instantiate a natural kind trope; or objects with some functional properties (chairs) instantiate some functional kind property (chairhood) or instantiate a functional kind trope; objects with some superficial properties instantiate some superficial kind property or instantiate a superficial kind trope. Aspects represent such superficial kind

properties or superficial tropes[73]. They thus represent an object as pertaining to a group of *superficially* similar objects, which group is determined by similarity in superficial surface body properties.

The continuous nature of the accuracy conditions of aspects is not only relative to some expected standard, it is also continuous because objective similarity relations in the world are themselves continuous. There are objects that are more typical of some superficial kind than others. A robin is a more typical bird than an ostrich. Perceiving a robin as having a bird aspect might thus be taken to be more accurate than representing an ostrich as having the same bird aspect. This might be translated by having representations with different levels of confidence (Morrison, 2016): one is more confident in representing the robin as having a bird aspect than representing the ostrich as having the same aspect. This level of confidence might be itself translated into specific (meta)perceptual experiences: the robin phenomenologically looks *more* like having a bird-form than an ostrich.

Ultimately, it might be that the continuity of accuracy conditions is not limited to aspects. Maybe *all* perceptual representations have such relative, continuous accuracy conditions. After all, it seems more accurate to represent a red patch as orange than to represent it as blue; or to represent a cube as a rectangular parallelepiped than as a spiky sphere. It is plausible (though I do not wish to commit on this point) that no perception is *fully* accurate due to noisy perceptual processes and limited cognitive resources. As interesting as these issues might be though, I will leave them aside, talking as if perception represents aspects in a dichotomous accurate or inaccurate way. This choice does not change the general idea that perception represents proprietary kind contents with objective accuracy conditions.

This idea that higher-level perception represents its own proprietary kind contents (distinguished from cognition) takes inspiration from an idea already expressed by Jack Lyons, who opposed most Liberals in claiming that, while perception was not limited to representations of low-level contents, it still didn't reach as high as representing natural kind or functional kind contents. Instead, Lyons argued that perception could only represent *perceptual* kind properties, which he defined through the notion of similarity, much like in the spirit of aspects:

> If perceptual kinds are genuine kinds in this sense, then they will have to exhibit the homogeneity and distinctness typical of kinds more generally, *and obviously the relevant dimensions of homogeneity and distinctness in this case will be perceptual*. Thus,

[73] It might be that another metaphysics is more suitable than a trope metaphysics for kind properties instantiations. I think it is quite useful and clarificatory to think of kind properties as being some *abstract* particular that is instantiated in objects as tropes. The natural kind property "cathood-trope" is a particular instantiation of the natural kind cathood. And the superficial kind property of "catform-trope" is a particular instantiation of the superficial kind "catform". Aspects attribute such superficial tropes to objects. These representational attributions can be more or less accurate, because superficial kinds (unlike natural kinds, but maybe like functional kinds) have no discrete boundaries, but instead have more or less typical instantiations relative to some standard.

> K is a perceptual kind just in case the members of K are perceptually similar to each other in some respects and perceptually different from other things in those same respects. To say that cows constitute a visual kind is to say both that cows look similar to each other and that cows have a distinctive look; i.e., they look unlike other things. (Lyons, 2005, p. 191, *my italics*)

Lyons argues that typical natural kind properties like "water" or "pine trees" do not form perceptual kinds, because objects having these properties are either not enough homogenous (they are too perceptually dissimilar between each other), or not enough distinctive (things from other kinds perceptually look like them). In the vocabulary developed in this thesis, we could say that natural kinds cannot form distinctive manifolds in a structured perceptual space. Lyons concludes:

> [T]here seems to be no systematic or otherwise interesting connection between perceptual kinds and the various types of natural kinds. Perceptual kinds have their own dimensions of homogeneity and distinctness, which are different from those of the various other sorts of kinds. (Lyons, 2005, p. 199)

My own defense of Liberalism endorses Lyons' point that one should distinguish between the high-level kind properties representable by perception (schematization) and the high-level kind properties representable by cognition (or any post-perceptual process). But Lyons' appellation of "perceptual kind" is I think slightly misleading. It suggests that high-level perceptual representations are about perceptual representations themselves, that they are higher-order representations having as content a property of perceptual representations. I have rejected such meta-representational view for perceptual Liberalism. The aspect Liberalism which I defend maintains that high-level perceptual representations represent genuine, objective worldly properties: *superficial* kind properties[74]. The notions of "perceptual kind" threatens to blur the distinction between the representational mental state (aspect) and its representational content (superficial kind properties or superficial kind tropes). The mental state is a "perceptual kind" in the sense that it relies solely on the manipulation of perceptual states. But what's represented by such a mental state is a genuine property of the world, not a meta-representational property.

To illustrate: one must distinguish between the biological kind "cow-*hood*" and the superficial kind of "cow-*form*". Here I use the term "form" as shorthand for talking about a group of particulars that is similar with respect to some superficial body properties such as shape, size, or texture. "Cow-form" is an objective kind property (categorical or relational-dispositional) of some particulars which would persist even in a world where no perceivers exist to perceive them. A superficial kind property of a particular is an instantiation of some

[74] My Liberal view would get closer to that of Lyons if I were to adopt a *relational* understanding of perceivable properties. Maybe it's not that aspects represent mind-independent, intrinsic, categorical (similarity) properties of objects, but they represent mind-dependent, extrinsic, relational (similarity) properties of objects. Still, I would insist that these properties remain objective properties that respond to norms of representational accuracy (but maybe Lyons would concur with this point).

form (just like a natural kind property of a particular is an instantiation of some natural essence).

Of course, particulars that instantiate the natural kind property of "cow-hood" strongly overlap with particulars that instantiate the superficial kind property of "cow-form". Simply put, biological cows often instantiate cow-form, and things instantiate cow-form are most often cows (instantiate cow-hood). This correlation is not necessary though. A cow might not have a cow-form if one has disguised it, if it is from a very unusual looking species of cow, or if we consider a cow fetus. Oppositely, things instantiating cow-form might not be cows. This might be the case for cow statues, cow drawings, cow holograms, or something disguised as a cow.

Parallelly, one should distinguish between the non-observational natural kind concept COW and the superficial kind representation cow-aspect. The former is a cognitive representation, a concept, representing a natural kind (cow-hood). The latter is a perceptual representation, an aspect, representing a superficial kind (cow-form). One could also token a purely observational concept of COW which represents merely perceivable superficial properties of cows, and might thus also be only referring to some cow-form property. Still, observational concepts aren't aspects. They are different representations, because observational concepts are still concepts – sub-propositional representations which can be tokened offline, non-demonstratively, and whose content depend on particular contexts of use –, while aspects are perceptual representations – nonpropositional representations which can only be tokened online, demonstratively, and whose content depends entirely on input stimulations and perceptual processes (schematization).

While we can refer to a high-level property that is cow-related in perception through aspects, this property doesn't correspond to the natural kind property of *cow-hood*. It corresponds instead to the superficial kind property of *cow-form*. The fact that aspects refer to superficial kind properties provides them with accuracy conditions that depend on some objective property of the world: surface body similarity (though these accuracy conditions are continuous and relative to some standard).

The fact that aspects represent *objective* superficial kind properties allows us to avoid a noxious issue: that of aspects having fully idiosyncratic and thus incommensurable contents. If aspects were only meta-cognitively representing a meta-representational fact about one's mental activity (that one's current perceptual representation is similar to some tokened memory representations), then no two individuals could have the same aspect representation when encountering the same particular. In fact, no two individuals have the same perceptual memory traces. Each and every individual has different perceptual stories. Furthermore, due to idiosyncratic learning histories, no two individuals will perceptually structure the same stimuli in equivalent ways.

Let's imagine Mary, from Canada, and Johnny, from Australia. Mary's perceptual memory traces and prototype representations of cows might be very different from Johnny's memories and prototype of cows. It is probable that Mary and Johnny have not seen *any* common exemplar of cow. If now Mary and Johnny are put in front of the same

cow in Italy (one neither has seen before), two quite different schematization processes will occur. In fact, the perceptual mnemonic traces that will be primed will be different for Mary and Johnny. Furthermore, the structuring processes underlying such similarity primings might also be different: Mary, an expert cow farmer, might automatically structure her low-level representation sub-personally selecting some shape-texture feature values in a very fine-grained way; while Johnny, an urban dweller, might structure his low-level representation by sub-personally selecting a shape feature value in a much coarser way.

Thus, because no two individuals have the same experiences, no schematization processes, and not aspect representations, are expected to be the same. This extends to different schematizations of one individual at different times of his life. Mary's schematization process for cows is different when she is 10 years old, compared to when she starts her farm, and when she is an accomplished cow farmer. Her structuring processes evolve through learning, and the memory traces that can be perceptually primed change through the accumulation and forgetting of perceptual memories.

If aspects were meta-representational, then we should conclude that Mary's and Johnny's aspect representation when they see the same cow is (type) different. Mary associates her current perception of the cow to similar idiosyncratic memory traces $\{\underline{M_1}, \underline{M_2}, \underline{M_3}\}$ in specific respects $\{R_1, R_2, R_3\}$, while Johnny associates his current perception of the cow to some other similar memory traces $\{\underline{M_4}, \underline{M_5}, \underline{M_6}\}$ in other respects $\{R_4, R_5, R_6\}$. We thus cannot conclude from this view that Mary and Johnny token the same cow-aspect. Instead, they would token fully idiosyncratic aspects: a $\underline{\text{cow-aspect}}_{\text{Mary}}$ and a $\underline{\text{cow-aspect}}_{\text{Johnny}}$. We should thus conclude ultimately that no two individuals can have the same aspect representation for the same particular.

This conclusion is problematic though. It seems more intuitive to claim that Mary and Johnny represent the same unique cow aspect type. If you ask Mary what she's seeing, she says "I see a cow", and Johnny might just reply "me too". It doesn't seem there is any incommensurability in their report of what they are seeing[75].

Furthermore, aspects should explain some complex behaviors (see IV.3.2 SCHEMATIZATION AVOIDS OVERINTELLECTUALIZING BEHAVIOR). They explain how an

[75] A difficulty here is that it is unclear, when Mary and Johnny report about what they see, whether they are reporting about aspects (seeing-as) or whether they are reporting about what they can *judge* from what they see (seeing-that). It is uncontroversial that Mary and Johnny might both possess the same COW concept. But we also have the intuition that Mary and Johnny would have the same *aspect* representation as of something having a cow-aspect. We can imagine that Mary and Johnny are skillful philosophers of mind, who are aware of the difference between perceptual representations proper and perceptual judgments (of course, we imagine that Mary and Johnny reject that perceptual judgments are part of perception). Now, we ask Mary and Johnny to imagine that they are conceptless individuals incapable of perceptual judgments. If Mary and Johnny accept the idea that we perceive aspects, I think we still get the intuition that they would be representing the *same* cow-aspect (attribute the same cow-form superficial kind), and that that's this aspect they would report by saying that they are both "seeing a cow".

organism is able to generalize one unique behavioral repertoire to different particulars. It is because we token the *same* aspect representation for these different particulars, and thus represent them as instantiating the *same* superficial kind property, that we can associate a unique behavioral repertoire for these particulars. Our behavioral repertoire toward one class of particulars might stay the same while our schematization process changes. This indicates that the aspect represented might stay the same while schematization changes. But this idea can only make sense if we recognize that aspects are not about other mental representations, but that they are about the world. Schematization can change while the superficial kind property that is represented remains the same. Mary at 10 might schematize cows in a way that is different from the way she schematizes cows at 45 when she has become an experienced farmer. Still, she represents the same worldly property: that of having a specific superficial kind property. that of having a cow-form. Likewise, Mary and Johnny might schematize the same cow in different ways, but still represent the same property of instantiating a cow-form. They can represent the same property because, as I argued, superficial kind properties are objective properties. Superficial similarity is a property of the world.

Aspects rationally guide behavior because they relate the perceiving organism to such unique property of the world. It is rational to potentiate the same behavior for objects that instantiate a common form, because objects with the same form are most often objects one can rationally behave similarly with. We come back here to the idea that superficial kind properties are expected to correlate with some other kind properties like natural or functional kind properties. In most ecological situations, something that instantiates apple-form also instantiates apple-hood. Ultimately, it is the apple-hood property that interests a behaving organism. It is this natural kind property that makes the apple nourishing, not its apple-form. But since apple-form correlates with apple-hood, it is rational for the organism to behave in a specific applehood-way when perceiving apple-form.

"Apple"-form, or 🍎-form[76], is not apple-hood. Plastic apples, apple holograms, and twin-apples do have the superficial kind property instantiating 🍎-form, but not the natural kind property instantiating apple-hood. Associating a specific behavioral repertoire with 🍎-form representations (🍎) thus stands merely as an ecologically valuable, economical, and rational association. In most situations, representing 🍎-form is useful as a proxy for apple-hood. This is obviously especially so for organisms who cannot represent apple-hood as such since this requires concept possession. A monkey thus can behave in the same way when representing 🍎, allowing it to have a generalized behavior when encountering different instances of apples. This monkey can act without conceptual thinking, and only through aspect perceiving. The fact that many organisms display such behavioral

[76] Talking of "apple"-form is misleading, because we are not necessarily talking of the form that apples have (we are not talking of apple-hood-form). Apple-form is a superficial kind subsuming apples, plastic apples, and twin-apples. That is why writing of the superficial kind 🍎-form and their representation through 🍎 is more rigorous and less misleading (though textually less fluent).

generalizations without mastering concepts speaks in favor of the existence of aspect representations (IV.3.2 SCHEMATIZATION AVOIDS OVERINTELLECTUALIZING BEHAVIOR).

The representation of superficial kind properties through aspects is not present in "conceptless" organisms. Human adults also represent such properties even though they possess corresponding concepts. One's mastery of the concept APPLE does not make one stop representing 🍎. But why should we persist in representing superficial kind properties through aspects when we have come to master corresponding natural or functional kind concepts?

The answer precisely lies in the *perceptual* nature of the aspect representations that represent such properties. Aspect representations, grounded in automatic, bottom-up schematization processes, are mentally fast and easy to produce. Conceptual categorization, on the other hand, requires more cognitive effort (access to semantic stores). Aspect representations in perception can come as a facilitatory step for cognition. One represents a particular aspect (🍎) and can learn to associate a particular concept (APPLE) with it. One can thus quickly and easily judge that some object is an APPLE simply by looking at it and schematizing it under a specific aspect.

Most importantly, in many situations, aspect representations are enough to act. One can efficiently behave in her environment by merely representing superficial kind properties. Something that instantiates 🍎-form can be grasped and eaten. It is to be expected then that even concept-possessors like human adults will many times behave merely based upon what aspects they perceive. I see Jack's face's aspect in the street (Jack is an old friend of mine). I thus smile and come closer. I do not need to think 'this is JACK' to do so. I simply act without thinking, based on what I *see*.

Of course, a concept-possessor might forestall the association of aspect with acts because she can believe that things are not what they perceptually seem to be. A human adult can enter a room and believe that the apples in the bowl are plastic reproductions. She thus does not come closer to eat them, even if she is starving. In other words, she dissociates her apple-aspect representation as of an 🍎-form (which is accurate) from her conceptual thought referring to apple-hood (plastic apples do not instantiate apple-hood). Contrarily, a starving horse entering the room will rush on the plastic apples and try to eat them. Only after a negative association has been put in place (it cannot eat it) will it stop trying to eat it. Before that, the horse's behavior is slavishly controlled by the apple-aspects it is perceiving.

V.3.4. ASPECTS AND MODES OF REPRESENTATION

Aspects represent superficial kind properties. Representing such properties allows organisms to behave in complex and rational ways with objects they have previously never seen. Different organisms can represent the *same* superficial kind property (e.g., *cow-form*, *apple-form, Joe Biden-form*) in different ways, through different schematization processes.

Mary and Johnny both represent the same cow-form, though through different schematization processes, i.e., through different structuring and similarity primings.

Aspect representations should primarily be individuated by the objective property which they represent, not by the *way* they represent it[77]. This is a classical idea in the philosophy of mind and language: we should distinguish between the property represented and the *mode of representation* through which we represent such property. It is commonly accepted that different modes of representation can represent the same property or individual. "Clark Kent" and "Superman" are words that represent the same singular individual in different ways (likewise for the CLARK KENT and SUPERMAN concepts).

This argument is easily extended to perceptual representations (Burge, 2022; Thompson, 2007b, 2010): it is empirically almost certain that my own perceptual experience ("qualia") of $\underline{\text{green}}_{24}^{\star}$ (a particular shade of green) is different from yours (maybe my own $\underline{\text{green}}_{24}^{\star}$ phenomenology is slightly lighter than yours). Even more dramatically, color inversion scenarios with fixed representational content are widely held to be conceivable, if not empirically demonstrable (Nida-Rümelin, 1996). Maybe my $\underline{\text{green}}_{24}^{\star}$ phenomenology is *green** while yours is *red**, though we represent the same $\underline{\text{green}}_{24}$ content.

Aspects also have this characteristic: they can represent the same property (a cow-form superficial kind property) through different *modes of representation*. These different modes of representation for aspects might correspond to different *schematization* processes[78]. Mary and Johnny schematize differently the same superficial kind property. Different modes of representation might also be actualized for the same aspect representation of a superficial kind property if such representations are articulated to different low-level representations. Just like seeing the same blue shade representation might be articulated to many varying representations of ambient illumination (thus yielding different experienced "perspectives", of modes of representation, of the blue shade), the same apple aspect

[77] One should be *very* careful with terminology here. I take aspects to be specific perceptual representations that attribute to objects a superficial kind property (i.e., classify them as pertaining to a superficial kind). I thus do not take aspects *to be* modes of (re)presentation. I am *not* saying that perceptual representations have a content (= "intentionality") and a mode of representation (= "aspectuality"), as sometimes put forth in the literature (Sacchi, 2011). I concur with the Fregean distinction, but my vocabulary differs. I think some (perceptual) contents are aspectual, and that such contents can come in many modes of representation. Notice that not just aspects have modes of representation on my view: a determinate low-level content of a shade of blue can come in many modes of representations too (e.g., when illuminated differently).

[78] I think the notion of modes of representation is best illustrated in perception through differences in phenomenal experience as in inverted spectrum scenarios. However, I do *not* think the notion of mode of representation in perception is reducible to modes of phenomenal experience. Instead, I think that phenomenal experience is grounded in perceptual modes of representation. I think that perceptual modes of representation are grounded in the perspectival nature of perceptual representations, and more precisely their dependence on other perceptual representations or sensory registration. Different schematizations are different modes of representation. This perspectival dependence is a representational property, not a phenomenological one. Even a phenomenal zombie could be said to have different perceptual modes of representation for the same property. Nonetheless, for ease of the argument, I talk of perceptual modes of representation through their phenomenal outcomes.

representation might be articulated to many varying low-level representations of ambient illumination, shape, color, texture etc. The general point is the following: varying modes of representation of the same superficial kind property depend on how the current percept is articulated to other representations, be they varying lower-level representations to which it is articulated, or varying perceptual memories to which it is associated through schematization.

How different modes of representation can get to have the same referent is a controversial and complex matter. It won't ultimately matter which solution we pick for aspects, as long as we recognize that different schematizations can represent the same superficial kind property. Let me say some words about some possibilities though.

One can argue reference is determined *satisfactionnally*: the referent of a representation is the object or the property that satisfies a particular description given by this representation. "The color of the sky" and "Albert's favorite color" are descriptions that might be satisfied by the same color kind property *blue*. Schematizations could be taken as some kind of perceptual descriptions: they describe a particular as being similar to other particulars in certain respects. Two different schematization descriptions might still be satisfied by the same superficial kind property. "Similar to $\underline{M_1}$ with respect R_1" might be satisfied by the same superficial kind property as "similar to $\underline{M_2}$ with respect R_2". Imagine $\underline{M_1}$ and $\underline{M_2}$ are both memory representations of cows, and that R_1 corresponds to head-shape feature while R_2 corresponds to trunk shape and texture. Still, the property that satisfies both descriptions might well be an instantiation of the cow-form kind.

The traditional opponent to such descriptive-satisfactional theories of reference are *causal* theories, which state that reference is determined by some causal relation of the representational mental state to its referent. My perceptual representation of $\underline{\text{blue}_{54}}$ refers to the property $blue_{54}$ because $\underline{\text{blue}_{54}}$ is (normally) caused by $blue_{54}$. Here again, different representations can come to represent the same property, because different representations can be causally bound to the same property (e.g., the same $blue_{54}$ property might have many different modes of representation depending on illumination conditions). Likewise, different schematizations could be causally bound to the same superficial kind property: cow-form instantiations might systematically cause a certain schematization for Mary, and another schematization for Johnny.

I won't commit to the descriptivist or causalist theory of reference here. In any case, both can agree on the following result: different schematization processes can represent the same superficial kind property (the same superficial form). So, Mary and Johnny's aspect representations might represent the same superficial kind property, even though their idiosyncratic schematization processes (structuring and similarity priming) are different. They might thus both be said to represent the same <u>cow aspect</u>, though through different modes of representation. To use again the superscript typographical convention to express modes of representation, we could write that Mary sees *o* as having a $\underline{\text{cow-aspect}}^{r1}$, while Johnny represents *o* as having a $\underline{\text{cow-aspect}}^{r2}$. The superscript expresses different modes of representation, i.e., different schematizations (or articulations to different lower-level

representations: Mary and Johnny might see the same cow-aspect having a different color, texture, orientation, size…). Like in the color case, these different modes of representation can be intuitively understood as corresponding to different phenomenal experiences (though not necessarily, see fn. 78 above): cow-aspect^{r1*} feels different from cow-aspect^{r2*} (just like green$_{24}{}^{green\star}$ feels different from green$_{24}{}^{red\star}$ in spectrum inversion scenarios, though they represent the same property).

This latter point is actually quite crucial to satisfactorily account for intuitions of phenomenal contrasts with high-level contents. Even though we recognize that one who can recognize pine trees has a different visual phenomenology compared to one who cannot, we do *not* have to recognize that all persons who can recognize pine trees have the *same* high-level visual phenomenology. This point was hard to make in previous defenses of Liberalism through phenomenal contrast arguments, since it was never made explicit which psychological mechanism was responsible for the high-level representation. Worst, if one were to think of such contrasts as explained by concept application, then one should conclude that all pine tree recognizers have the same high-level recognitional experience, since they all apply the same PINE TREE concept. Having to accept this conclusion would make phenomenal contrast argument much less persuasive.

My *aspect* Liberalism which integrates *modes of representation* avoids this conundrum. I claim that pine tree recognizers have the perceptual capacity to schematize pine trees in a specific way. They come to learn to represent pine tree aspects, and thus to perceive pine-tree-form. But because such schematization processes might widely diverge between different pine tree recognizers, there will exist many modes of representation for the same aspect representation, many *ways* to perceive the same pine-tree-form. One forester from Canada thus represents pine tree aspects^{r1}, and another forester from Italy represents pine tree aspects^{r2}. These modes of representation of the same aspect content can correspond to different phenomenal experiences. With aspect Liberalism, one can thus still hold a Liberal view from phenomenal contrast arguments while not having to conclude that all pine tree recognizers have a unique pine tree visual experience.

V.3.5. Externalist and Internalist Individuations of Aspects

A final worry might concern the *individuation* of aspect representations: even if we say that Mary's and Johnny's aspect representations represent the same property, is this enough though to conclude that they are the same representations, the same *cow*-aspect?

This ties to a broader philosophical debate about individuation conditions for mental representations. "Externalists" or "anti-individualists" tend to favor individuation of mental representations in terms of referred to properties and individuals, i.e., in terms of *extension*. When Joe and Jack talk about arthritis, their concept of ARTHRITIS is the *same* concept representation, because it represents the same thing (joint inflammation) even though Joe and Jack might actually believe very different things about arthritis (Burge, 1986). "Internalists" or "individualists" tend to favor individuation of mental representations

in terms of the very information that is carried out *for the individual* by such representation (be it true or false), in terms of computational-functional role, or in terms of the phenomenal character of such representation, i.e., in terms of *intension* (I remain intentionally vague here, since there are many ways one can understand internal, intensional individuation of representational states).

Such internal characteristics of a representational state can all be considered as some different *modes of representation* of the same property or individual. One and the same property can be associated with different information (e.g., a concept can point to different stores of information) and/or can have a different phenomenal character for different individuals. Internalists argue that representations should be individuated by such internal factors, *not* by the referred to property or individual. When Joe and Jack talk about arthritis, an internalist would say their concept of ARTHRITIS is not the same representation, since Joe and Jack believe radically different things about arthritis, thus they have different concepts of ARTHRITIS that have different cognitive significance *for them*. Likewise, an internalist would argue that perceptual states should be individuated by their specific internal perspectival character, corresponding to phenomenal character, not by their representational extension. A $\underline{\text{green}}_{24}{}^{green\star}$ experience and a $\underline{\text{green}}_{24}{}^{red\star}$ experience should be regarded as different perceptions, even though they might represent the same green shade as in color inversion scenarios.

Adjudicating between externalists and internalists theories of mental individuation is far beyond the scope of this thesis. I would notice here that at an intuitive level, both theories seem to be relevant for different explanatory purposes. We have both the intuition that Joe and Jack talk about the same thing (*arthritis-hood* understood as some medical kind property), but also that they don't (Joe talks about arthritis-for-Joe, and Jack about arthritis-for-Jack); we have the intuition that Joe and Jack see the same green property (even though they can have different phenomenal experience of such property), but also that they don't (Joe has a *green** visual experience, Jack as a *red** visual experience). Finally, we might have both the intuitions that Mary and Johnny see the same cow-aspect (since they represent the same superficial kind property), but also that they don't, since their schematizations grounding representation of this property are different (they might not emphasize the same perceptual feature of the cow and implicitly prime different memories of cows).

It is not absurd to be open-minded here, and to argue externalist or internalist individuations might serve different explanatory purposes. The externalist individuation for instance accounts for the possibility of communication and correction (because Joe and Jack talk about the same property). The internalist individuation on the other hand accounts better for individual-relative thought processes and experiences (Sacchi & Voltolini, 2017): it explains why Joe passes his medical exam and Jack doesn't (because Joe's information carried by his own ARTHRITIS concept is true, while Jack's is false); it explains why Joe and Jack do not have the same visual experiences.

We can keep the same open-minded attitude towards the individuation of aspect representations. For some explanatory purposes, it might be useful to individuate aspects in

terms of their extension, i.e., in terms of superficial kind properties. In that case, two persons having different schematizations for the same aspect would still be said to token the same (type) aspect. It would explain how Mary and Johnny coordinate their perceptually-based behavior. If Mary asks Johnny "go closer to the cow", Johnny can safely understand that Mary asks him to get closer to the thing with a cow form[79]. This is so even though Mary's and Johnny's schematizations, i.e., modes of representation, of such property are different. They should be said to represent the same property, and have the same aspect representation, if we want to understand how Mary and Johnny coordinate and communicate according to what they see. This is the approach I took in the previous section.

In other explanatory contexts though, it might be important to distinguish between the different schematizations of the same superficial kind property. Mary the expert farmer will not schematize cows in the same way as Johnny the urban dweller. Mary won't be attending to the same feature dimensions of the cow, and she won't be priming the same memories as Johnny. Thus, Mary will see a cow-aspect^{r1}, and Johnny a cow-aspect^{r2} (here I underline the modes of representation determined by schematization and make them part of the representational content). Furthermore, Mary is expected to not only schematize a cow-aspect^{r1}, but also to schematize more fine-grained sub-aspects, such as a Limousin cattle aspect^{s1}.

Thus, explaining perceptual experience might be a context in which individualistically individuating aspects (counting modes of representation as part of the content) is relevant, especially, of course, if one accepts a kind of Fregean representationalism for perceptual experience (Chalmers, 2006; Sacchi, 2018; Thompson, 2009), as I advocate. Perceptual phenomenology is determined by mental schematization, not by what this schematization represents (the superficial kind property). In this context, it might be reasonable to argue that Mary and Johnny token different, idiosyncratic aspect representations corresponding to different high-level phenomenal "aspectual" experiences.

Ecumenism might thus be warranted for aspect content individuation, just as for individuation of all other representational contents. "Content" might have two reasonable understandings, one extensional (corresponding to the superficial kind property represented), the other intensional (corresponding to the schematization process). Which understanding of "content" we choose might depend on explanatory context. Any representation thus has two compatible contents: one extension, and one intension (or mode of representation). Aspects represent fixed extensions through different modes of representation. Aspects represent fixed superficial kind properties through different schematization processes.

[79] Notice that Mary could ask the same question, and Johnny could have the same behavior, even though Mary is pointing at a *papier-mâché* cow. What both have in mind is the cow superficial kind property, not the cow biological kind property.

VI. KNOWING HOW TO SCHEMATIZE

At this point, one might legitimately wonder how the perceptual system "knows" how to schematize, i.e., how the perceptual system "knows" which features to build and select for structuring and for similarity priming. We do not want schematization to be constitutively dependent on a process of hypothesis testing of the following form: I wish to know whether '*o* is an X?', I know what are diagnostic features of X, therefore I schematize in a X-relevant way on such known feature dimensions. We (as Liberals) do not want it because that would mean schematization is constitutively dependent on rule-based cognition (hypothesis formation), which would annihilate the conclusion from the last chapter. A question we must answer is thus: how can the perceptual system *on its own* build and select relevant features for schematization?

The core of the answer lies in the capacity for perception to know how to schematize independently of cognition. We can distinguish between three ways perception comes to know how to schematize: phylogenetic selection, ontogenetic perceptual learning, and synchronic attentional modulations. I will say a few words about each of them in turn, showing that none of these processes depends *constitutively* on cognition. Still, they can surely be *influenced by* cognition in a causal way. Schematization can thus be *causally* dependent on cognition. I show that this causal dependency is contingent (not necessary). I further argue that accepting such top-down cognitive modulations of schematization is not equivalent to accepting *cognitive penetration*.

In this final chapter, I will first talk of forms of schematization learning that are strictly independent from top-down influences of cognition (VI.1), including evolutionary species-scale learning and perceptual individual-scale learning. Then, I will consider how schematization can be voluntarily controlled through top-down cognitive modulations (VI.2), ending up with considerations as to whether such modulations should be taken as cases of cognitive penetration.

VI.1. AUTOMATION

VI.1.1. EVOLUTIONARILY PREWIRED STRUCTURING

The first way we come to "know how to schematize" is through a process of evolutionary selection. There are certain low-level features which automatically drive perceptual

structuring because these are features that yield adaptively valuable similarity primings and aspect representations.

A topical illustration of this idea is provided by a study by Joshua New, Leda Cosmides and John Tooby (New et al., 2007). These authors put in place a classic change detection task, wherein two very similar pictures alternate, except one picture is a slight variation of the other (e.g., one object disappears in one picture). The goal in a change detection task is simply for subjects to detect the changing part of the picture. Change detection tasks are made quite hard by placing brief blank fields in-between alternating images (Rensink et al., 1997). This result is usually interpreted as showing that attention to a certain place or object is required to detect change. Placing a blank field in-between presentations resets such attention and thus strongly disturbs change detection. The idea of Joshua New and collaborators' study was to test whether change detection is faster for certain evolutionary important stimulus categories, especially the categories "human" and "animal". The hypothesis behind such a test is that "humans" and "animals" might be stronger attractors for attention than artificial objects such as plants, houses, or tools, because humans and animals are ancestrally important things for humans to detect. The results of New et al.'s change detection study confirmed this hypothesis: participants were faster at detecting changes when these concerned humans or animals in the picture, as compared to plants, tools or fixed landmark objects. Thus, the authors conclude that:

> Taken together, the results herein implicate a visual monitoring system equipped with ancestrally derived animal-specific selection criteria. This domain-specific subsystem within visual attention appears well designed for solving an ancient adaptive problem: detecting the presence of human and non-human animals and monitoring them for changes in their state and location. (New et al., 2007, p. 16603)

The attentional grasping of animal stimuli shows that some kind of animal classification has already been carried out pre-attentionally. Remains to be determined whether it corresponds to pre-attentional schematization (animal-aspect) or categorization (ANIMAL).

We have good empirical reasons to favor the idea that attention is attracted to some unconsciously perceived diagnostic low-level features and is not the result of some categorization proper. In a study, Bria Long, Viola S. Störmer and George A. Alvarez (2017) demonstrated a significant reaction time advantage when one is searching for an animate stimuli among inanimate distractors (and *vice versa*). Experimenters explicitly tested whether such reaction time advantage was due to the representation of categorical differences (ANIMATE vs. INANIMATE objects), or whether it was due to attention being attracted to some specific animal-diagnostic features of the scene, with "curvature level" being the central suspect. Long et al. showed that the reaction time advantage was actually preserved even if the visual search task's target and distractors are replaced by "textform" transformations of animals/nonanimals which preserve curvature information but become utterly uncategorizable by subjects. This directly shows that attention is not grasped by high-level categorization into ANIMAL, but it is grasped by certain low-level features (here,

curvature) that are diagnostic of animate/inanimate stimuli. In my framework, I would say that such textform transformations are still perceived as having an animate-aspect or inanimate-aspect. Curvature information is enough information for structuring stimuli in an animal-relevant way and priming distinctive memory traces of animals. Of course, because such stimuli are either not attended or present highly deteriorated information (textforms), such schematization might be carried out with a low level of confidence. One sees an animal-aspect but is unsure about it (that might be enough though to complete animal-detection tasks above chance). Once attention has been attracted on the stimulus, input information is made more precise, and the same schematization as of an animal aspect might be carried out with much more metaperceptual confidence.

Because the detection of other animals is adaptively crucial, it is quite unsurprising that evolution would have hardwired an automatic schematization process along specific curvature feature values if these curvature values normally correspond to that of animals. Automatically structuring perceptual representations around curvature feature values allows the fast perceptual discrimination of animal-aspects, which might be a sufficient heuristics to produce rational behaviors (first of all: attentional actions to confirm or reject the first impression as of an animal aspect).

This idea of an evolutionary hardwiring of some schematization processes is even more convincing if we consider the fact that it is adaptively crucial that animal-detection should be very fast and shouldn't depend on complex, cognitively costly inferential processes. This is exactly what an evolutionary automatization of schematization gets us: it adaptively *offloads* to perception (K. Connolly, 2019; Machery, 2016) the capacity to represent an important information, i.e., that something that instantiates animal-form is present (by representing an animal-aspect). Through such perceptual offloading, humans can bypass the need to appeal to cognitive, conceptual resources to detect animals.

It remains an open empirical question to determine what feature dimensions are automatically selected in structuring, and whether there really are some of these structuring processes that are evolutionarily hardwired from birth. Besides *curvature* for animacy detection, another good candidate that has attracted empirical considerations is "Fourier amplitude information"[80] for the detection of faces. Faces are demonstrated to be detectable at ultra-fast speed and yield pop-out effects like animals. Specific amplitude information is thought to be highly diagnostic of faces (Crouzet & Thorpe, 2011; VanRullen, 2006).

One could argue that evolution has hardwired specific selectivity to this low-level feature so that when we perceptually structure a perceptual input, we automatically select for specific amplitude values, and we thus prime representations of already-seen faces with similar amplitude values, thus yielding a primitive face-aspect representation. Of course, in

[80] "Fourier amplitude information" is a statistical spectral property of *images*, relating roughly to "the global distribution of orientations and spatial frequencies" (VanRullen, 2006, p. 3023). Mathematical details do not hinge on the argument. Notice that one does not need to maintain that we *perceive* amplitude information. The point is only that the visual system is sensitive to such information in its low-level representations.

most ecological situations, we have access to much more information than low-level amplitude values of scenes. But some primitive face aspect representation along this simple feature dimension might be a relevant suboptimal fast-and-frugal schematization heuristic for the detection of faces. Again, this might just be a first schematization step with low perceptual confidence which serves to invite further cognitive processing (e.g., by attracting attention on the detected object with a face aspect).

Whatever the low-level properties chosen, the idea that there must be some specific low-level dimensions to which evolutionary hardwired bottom-up perceptual structuring mechanisms automatically attunes us to is strengthened by empirical findings showing that infants as young as 2-months old (Kosakowski et al., 2022), and maybe even human fetuses[81] (Reid et al., 2017) spontaneously attend to face-like stimuli. This tends to show that at least some structuring processes are phylogenetically hardwired. As soon as some objectual mnemonic traces of faces are encoded and retrievable by the infant (through similarity priming), infants can come to represent face aspects, and their attention is grasped by such behaviorally crucial representation. Saying that infants perceive face aspects is a better account of their attentional behavior than saying that infants are merely attracted to some low-level feature like certain amplitude information values (*why* would they be interested in such feature?), or to say that infants are born with an innate concept of FACE.

Nonetheless, saying that schematization, and thus aspect perception, is *hardwired* or *innate* is probably exaggerated. After all, babies are not born with hardwired perceptual memories[82]. What babies might be hardwired with are specific bottom-up structuring processes, like automatically filtering for some amplitude information values diagnostic of faces. At first, these specifically structured representations do not prime any memories. But since babies spend a lot of time staring at caregiver faces, they soon come to store many perceptual traces of faces. Similarity priming, and perception of face aspects, will thus soon follow. We might thus more prudently say that babies are born with some *prewired* structuring mechanisms which allow to quickly bootstrap schematization as of face aspects. It is adaptively useful for babies to come equipped with such prewired structuring process, since detecting face aspects is one of the most important perceptual skill that an infant should master early. If the infant needed to learn "from scratch" which features are relevant to filter and attend to for face schematization, the process would take longer to learn and be less adaptive. Thus, some structuring processes might be at least evolutionarily prewired for bootstrapping important schematizations.

---

[81] Fetuses are perceptually sensitive to basic visual stimuli projected through the uterine wall. It is thus demonstrable that fetuses, like newborn babies, are more likely to engage (head turn) towards upright face-like configurations than equivalent upside-down stimuli (Reid et al., 2017).

[82] A fascinating possibility could also be that some kind of prenatal fetal memory is enough for bootstrapping schematization at birth. Fetal memory might be acquired via intrauterine experiences, and maybe through spontaneous simulations and fetal dreaming (Hepper, 1996). Maybe fetuses already dream about "face"-like things. Newborn babies might thus already store memory traces of face-like "representations", which allow them to schematize face aspects at birth.

VI.1.2. Diachronic Perceptual Learning

The number of superficial kind properties that can be perceptually represented from birth through prewired schematization processes is probably very limited, and only concerns aspects that are universally adaptive for social animals like humans (e.g., face-aspects, animals-aspects, body-aspects).

But beyond these phylogenetically basic perceptual aspects, there must be many more aspect representations that are *learned* during the lifetime of the subject. A strong prima facie argument for this idea is the existence of perceptual experts (Chudnoff, 2020; Stokes, 2021a, 2021b). Some experts seem to be able to perceive properties that novices cannot. Typical examples include the perception of subordinate bird species by ornithologists, the perception of tumors in X-ray images by radiologists, or the perception of specific chess positions by professional chess players. Discussions about perceptual expertise are intimately related to the debate over the admissible contents of perception. In fact, interpretive contentions arise concerning whether such experts should be viewed has having acquired a genuinely unique *perceptual* recognitional capacity that is perceptually unavailable to the novice (e.g., recognizing "Garden warbler", "lung tumor", "Catalan Opening"), or whether alternative Conservative explanations should be considered. These alternatives should sound familiar: one might claim that perceptual experts are actually *cognitive* experts (experts in "intellectual impressions" but not experts in forming "sensory impressions" to take Elijah Chudnoff's classification). Or one might also claim that perceptual experts merely learn to represent new *low-level* properties that the novice does not see (Dubova & Goldstone, 2021; Goldstone, 1994a; Schyns et al., 1998). These learned low-level contents are then exploitable for further categorical inferences or causing new intellectual impressions or intuitions.

Those who defend that perceptual expertise is genuinely *perceptual* also frequently appeal to common hallmark arguments (Stokes, 2021b). Obviously, such appeals will suffer from the same difficulties that I have described in the ambit of the admissible contents of perception debate (see II.6. Limits To Common Hallmark Arguments). In the ambit of my own approach, I have focused on finding a plausible mechanism for high-level representations in perception, which led me to uphold schematization as the best candidate.

The relevant question for perceptual expertise thus becomes: is there anything like expert schematization? More generally: can we learn to schematize in new ways during our lifetime? Perceptual "expertise", if it exists, would illustrate a more general idea, which is that we can learn to perceive new superficial kind properties in our environment through perceptual learning. Many of us are also "experts" in representing car aspects, cat aspects, or Donald Trump aspects. While the term "expertise" is usually reserved for exceptional and rare capacities compared to an average perceptual standard, what interests us here is the more general capacity to perceptually *learn* to schematize beyond the database of perceptual schematizations that are phylogenetically prewired.

So, is there a plausible story to tell about *learning to schematize*? I think the answer is positive. Schematization is defined by conjoined structuring and similarity priming of perceptual memories. We come to learn new memories by simply having new experiences of specific objects. Perceptual experts store *more* perceptual memories of some particular thing than novices. Augmenting the store of memory exemplars (or refining some prototype memory) eases subsequent similarity priming. Furthermore, experience with some particular domain might create some specific priming anticipations through classical association learning. If an aspect statistically co-occurs often with other aspects, then such aspects might come to be associated through classical association learning. For instance, one might come to associate her daily office aspect with her colleague's Mary's face aspect. Upon perceiving the office aspect, one might thus come to implicitly prime perceptual memories of Mary's face. This eases the schematization of Mary's face aspect upon seeing Mary. Notice that such association is *not* assimilable to *similarity* priming. Similarity priming in schematization requires the presence of a (structured) perceptual representation that primes similar memory traces *based on representational similarity*. Classical association learning relies on associations governed by some statistical co-occurrence. Only similarity priming and schematization ground aspect representation. Statistical association does not (it's not that upon seeing one's office aspect one come to perceive Mary's face aspect. Trivially, representing the latter aspect requires that a face instantiating Mary's face-form be present. But association learning can ease subsequent similarity priming[83]).

Besides improvements of perceptual memory and similarity priming, the perceptual learning literature clearly demonstrates that we can learn to automatically structure and attend differentially to objects (K. Connolly, 2019; Goldstone & Byrge, 2015; Ransom, 2020a). As psychologists Linda Smith and Diana Heise have clearly put it, "[p]erceptual similarity – perceptual structure – is dynamic because perceptual structure is not a thing; it is the emergent result of attentional processes" (L. B. Smith & Heise, 1992, p. 267)[84]. Schematizations are not phylogenetically fixed mental processes (excepting some prewiring of them). Schematizations also dynamically adapt to environmental pressures in one's lifetime. Again, perceptual experts are good illustrations of this idea.

If the mycologist wants to differentiate between golden chanterelle aspects and jack-o-lantern aspects, she will need to attend to very specific feature dimensions so that both mushroom representations are structured in different ways in high-level perception and are

[83] Besides such priming anticipations through association learning, one might think that the perceptual system is constantly producing perceptual anticipations of what is going to be perceived (including aspects), if one accepts something like a predictive processing view of perception (Hohwy, 2013). In that case, we could accept that similarity primings are *systematically* anticipated, and that such anticipations can be improved through experience (because our "priors" or "models" of the world improve). As interesting as such view might be, I won't consider it further here.

[84] Structuring is more general than attention in my vocabulary, though attention is clearly a strong modulator of the structuring process (by modulating which features are emphasized or how they are organized into holistic configurations).

organized in distinctive representational manifolds. Because golden chanterelles and jack-o-lanterns are structured differentially by the mycologist expert, they will prime differentiated perceptual memory traces. An expert mycologist thus sees golden chanterelles and jack-o-lanterns under different golden chanterelle aspect and jack-o-lantern aspect. The novice on the other hand does not possess the learned automatic bottom-up attentional skills of the expert. Golden chanterelles and jack-o-lanterns are structured indiscriminately in high-level perception and correspond to tangled, undifferentiated manifolds[85]. Thus, golden chanterelles and jack-o-lanterns prime indistinct perceptual memory traces for the novice, and she thus at most sees an indiscriminate mushroom-aspect.

To repeat an important point made in V.2.2 NO LUCK IN ASPECT PERCEPTION, note that *even if* the novice, by sheer attentional luck, structured in perception a golden chanterelle just like the expert, it won't be enough for her to represent a differentiated golden chanterelle aspect. In fact, it is expected that the novice won't prime distinctive perceptual memory traces of golden chanterelles. The novice will simply prime indiscriminate memories of mushrooms. This is so because the memories of mushroom stored by the novice are themselves structured in a coarse-grained way. The novice does not have memory traces of golden chanterelles which only keep with golden chanterelle relevant features. The novice would have to keep in memory relevantly structured representations of golden chanterelles to prime these memories specifically when she encounters again a golden chanterelle. This of course won't happen if the novice structures in a relevant golden chanterelle way only by sheer luck. It will happen only if the novice consistently (sub-personally, bottom-up) applies a specific perceptual structuring process to golden chanterelles, consistently stores perceptual memories of these representations, and discriminately primes these memories when re-encountering a golden chanterelle. In other words, it requires that the novice learns a perceptual schematization skill.

Bottom-up attention plays a central explanatory role in many psychological and brain models of perceptual learning (Maniglia & Seitz, 2018; Roelfsema et al., 2010). For instance, in the "attention-gated reinforcement learning" model of Pieter R. Roelfsema and collaborators, perceptual learning depends on the correspondence between attention to a specific feature and the release of reward neuromodulators (dopamine and acetylcholine). Such neuromodulators are known to increase neural plasticity, explaining why their release when attending to a specific stimulus feature strengthens neural connectivity for this feature representation. Oppositely, when reward modulators are released, non-attended feature representations will tend to have inhibited connectivity. The attention-gated reinforcement learning scheme explains how certain patterns of (bottom-up, endogenous) attention are strengthened, while others are inhibited in perceptual learning. "Reward" here

[85] Of course, this doesn't mean that golden chanterelles and jack-o-lanterns are perceptually indiscriminate for the novice. They are still different regarding their low-level contents. It's just that they are indiscriminate regarding the aspects under which they are both represented (i.e., as having a unique mushroom aspect).

can be understood as some kind of behavioral success (either in a controlled lab task or in natural environments).

But how does one select the relevant feature dimensions to drive perceptual structuring in the first place? We can distinguish two broad strategies following the perceptual learning literature (K. Connolly, 2019; Goldstone et al., 2011; Ransom, 2020a). One might first *randomly* attend to some feature dimensions, a process often coined as "blind flailing" (Goldstone et al., 2011). In a similar vein to natural selection mechanisms, whereby random genetic mutations are unintentionally passed over generations when they produce a successful phenotype, in blind flailing random attentional modulations are also selected and automatically reproduced when they produce successful (rewarded) behaviors. Blind flailing is thus a kind of *unsupervised* perceptual learning.

One can illustrate this idea with a scenario of meeting twins for the first time. At the beginning, one has much trouble discriminating between the two twins, but slowly one learns to distinguish them efficiently. This learning is not (necessarily) intentionally guided nor conscious: one does not need to know and be conscious of which feature one is attending to for the successful discrimination. The attentional pattern has simply been selected through attentional blind flailing and reinforcement selection (e.g., as described in the "attention-gated reinforcement learning" model of Roelfsema and collaborators). Blind flailing is a case of attentionally lucky discrimination successes that get stabilized and automatized through reinforcement processes. One thus comes to learn different schematizations for the two twins and comes to perceive different aspects for them.

The second way perceptual structuring can be modified is through *targeted* attentional patterns, or what Robert L. Goldstone and collaborators call "myopic flailing". Under myopic flailing, one intentionally attends to certain feature dimensions, because one has some beliefs about which dimensions are relevant to succeed in a particular recognition task. Myopic flailing is thus a kind of *supervised* perceptual learning.

For instance, one might be told by the twins' parents that the most distinctive physical trait between the twins is the shape of their nose. In that case, one will *myopically* flail her attention to the nose of the twins to succeed in recognizing them. If really successful for recognition (if the parents didn't lie), such attentional patterns will be reinforced and automatized. We can only talk of some kind of "myopic" flailing because our intentional attentional commands are too coarse-grained to be able to target specific feature dimensions of structuring in a sufficiently fine-grained way. "Attending to the nose" does not tell one which specific locations, curvature values, or subtle gestaltic relations should be driving structuring. As Robert Goldstone et al. say:

> Myopic flailing conveys that people can educate their perceptions more efficiently than expected via pure random variation, even though their manipulations are less direct and straightforward than they would be if they could access and manipulate all aspects of the perceptual module. (Goldstone et al., 2011, p. 8)

We can thus learn to automatize structuring processes in schematization through blind and myopic flailing processes that when successful get reinforced. Newly structured perceptual representations furthermore will refine the perceptual memory traces that are stored, thus refining the similarity priming process. A virtuous structuring-priming circle has been set up.

While I have insisted on perceptual learning as modifications of bottom-up attentional selection, its role in structuring is more extended than this. In fact, I have only considered a process whereby a *fixed* perceptual space serves as the basis from which an individual learns to filter relevant feature dimensions to carry on relevant schematizations. However, one might reasonably hold that we can perceptually learn to literally perceive *new features* (Schyns et al., 1998). One can come to learn to represent new low-level properties, thereby widening the schematization possibilities.

Two mechanisms can here be invoked again. One can learn to *differentiate* between feature dimensions that were previously confounded. By differentiating feature dimensions that were previously undifferentiated, perceptual learning can improve the structuring process. In fact, differentiation makes available new features to emphasize and combine, and thus new features fueling implicit similarity primings. One can thus learn to see a new, more fine-grained aspect that one could not see before.

One might for instance learn to perceptually differentiate between subtle low-level textural differences on the trunk of mushrooms, say the skin texture t_1 of golden chanterelles and the skin texture t_2 of jack-o-lanterns, whereas one was previously only seeing an indiscriminate skin texture t_0 for both mushrooms. Before learning to perceptually discriminate between t_1 and t_2, structuring on the trunk-texture dimension was structuring only with the indiscriminate t_0 value for golden chanterelles and jack-o-lanterns. One thus could only prime the *same* perceptual memories for golden chanterelles and jack-o-lanterns in respects to texture, and represent the same indiscriminate aspect for them. In other words, one was seeing golden chanterelles and jack-o-lanterns under the same unique mushroom aspect. But after learning to perceptually differentiate between t_1 and t_2 through low-level perceptual learning, structuring on the trunk-texture dimension can separate different manifolds that are separated in virtue of their different trunk texture values. These manifolds will prime differentiated memory traces: one can now see a mushroom as having distinguished golden chanterelle or jack-o-lantern aspects.

Opposing perceptual *differentiation*, one can also perceptually learn to represent a new feature by *unitizing* features that were previously parted. A classical illustration of such unitization is the automatic holistic or configurational processing of chess positions by professional chess players. A professional chess player might for example learn to perceptually unitize (perceive holistically) certain combinations of pawns and knights so as to become perceptually sensitive to Catalan-Opening-gestalts (Chase & Simon, 1973). Before such unitization, the novice chess player only represents an arrangement of individual pieces at individual locations. Schematization is then undifferentiated between several openings, and one might only be able to token an indiscriminate chess game aspect.

After unitization though, a professional chess player can perceptually represent configurational (low-level) properties like Catalan-Opening-gestalts, which stay constant even when irrelevant pieces are at different places. If the chess player emphasizes in high-level perception such specific configurational property, she will prime memory traces that specifically have the same Catalan opening configurational property. The chess player will thus see not only a chess game aspect, but also a Catalan opening aspect. Many cases of perceptual expertise are thought to rely on such holistic, gestaltic kind of representation. Face perception is another classic example: we might come to soon rely on such holistic features for face schematization (instead of relying on primitive features like amplitude information). Such holistic unitization has been observed for many other high-level contents, such as cars, dogs, or fingerprints. This specialized perceptual structuring (unitization) is obviously not innate but must depend on intense perceptual training. Learning to represent these holistic features explains how experts come to know how to perceptually represent (through schematization) distinguished car aspects, dog aspects, or fingerprint aspects.

While classical examples of perceptual expertise are good examples of how schematization processes can come to be automatized through perceptual learning, the range of aspects that come to be automatically represented is far from being limited to these cases. We all come to automatically represent idiosyncratic aspects in our daily environments through perceptual learning. This goes from aspects of many specific people (Mary aspect, Joe Biden aspect, Napoleon aspect), animals (bird aspect pigeon aspect), plants (mushroom aspect, golden chanterelle aspect), natural scenes (waterfall aspect, mountain aspect), or the paraphernalia of man-made daily life objects (hairdryer aspect, cellphone aspect). We also come to represent fully idiosyncratic subordinate aspects: my cat aspect, my sister's house aspect, my own hairdryer aspect etc. The point is that automation of schematization is not in the sole purview of perceptual experts, but it is a widespread and general psychological mechanism which is crucial in making perception useful in the guiding of our daily behaviors. Perceptual experts are special only in making this mechanism apply in highly specialized, rare, and socially valorized domains. But all of us are perceptual experts of their daily surroundings, having learned to automatize specific structuring and similarity priming processes (i.e., having automatized aspect representations) that are relevant in this environment.

VI.2. Control

VI.2.1. Synchronic Cognitive Modulations

The preceding two sections presented how schematization, and thus aspect perception, can come to be automatized through evolutionary prewiring or perceptual learning. These are processes whereby one acquires a kind of subpersonal, *procedural* knowledge as of how to

schematize. But another important way one knows how to schematize is through voluntary, synchronic control of the schematization process. We can change voluntarily, *on the fly*, the schematization process, and thus choose under which aspect we are perceiving an object. We already encountered a clear illustration of this voluntary control in V.1.1 THE NONCONCEPTUAL REPRESENTATION OF ASPECTS, when I presented how one can, through different schematizations, switch between seeing a rabbit aspect and a duck aspect in front of the duck-rabbit picture. I here repeat and generalize some of the ideas hinted at in this previous chapter.

As we saw V.1 WITTGENSTEINIAN VARIATIONS, we can distinguish between two competing interpretations of bistable effects. On the one hand, some philosophers argue that the switch in perception amounts to changing the concept that applies to the perceptual representation (Fodor, 2007; Prinz, 2006). The change between the duck-percept and the rabbit-percept is explained by the fact that in the first case we see the object as a DUCK and in the second case we see the object as a RABBIT. The issue with such position is that it might overintellectualize the phenomenon of perceptual bistability (MacPherson, 2006; Orlandi, 2011). It is unclear from the empirical literature whether bistability requires the possession of genuine concepts, or whether having memorized perceptual exemplars of the two percepts is sufficient (Wimmer et al., 2011). Furthermore, and most importantly, if one accepts a nonconceptual view of perceptual contents (as I have argued we should), then changing the concept that is applied to an object should not be viewed as a *perceptual* effect, but as a *cognitive* one. While clearly the perceptual *judgments* we make might switch with the percept, it doesn't explain the genuinely *perceptual* phenomenology of bistability, or the fact that the *image itself* seems to change when a switch happens, and not just our judgment about it.

A second popular explanation for figure bistability is to explain it in terms of switches in attentional patterns toward low-level properties (K. Connolly, 2014; Fish, 2013). When one is seeing the duck, one might be visually attending in a leftward biased way (i.e., towards the duck's beak), whereas when one is seeing the rabbit, one might be visually attending more in a rightward way (i.e., towards the rabbit's mouth). Under this explanation, bistability can only be a *low-level* bistability: the switch is explained by different low-level properties being made more or less salient by attentional modulations.

Experimental studies have demonstrated that top-down attentional control plays a central role in the percept-switches involved in ambiguous figures (Meng & Tong, 2004). If one accepts that attentional changes produce changes in perceptual phenomenology (Carrasco et al., 2004; Watzl, 2017), then this view naturally explains bistability. The view moreover happily avoids presupposing conceptual mastery as a necessary requirement for bistability. One only needs to know where and what to attend to get bistability.

Nonetheless, one might complain that the attentional view remains phenomenologically unsatisfactory. In fact, it doesn't seem as if, when we switch from the duck to the rabbit and *vice versa*, we are merely moving our attentional focus around. Rather, one usually talks of a full recognitional switch: the figure is now experienced as a

duck and then as a rabbit. Location-specific attention or feature-specific attention cannot account for such a recognitional change.

We can solve the tension between these two interpretations by appealing to the idea of a voluntary switch in perceptual schematization. Neither is the switch one in concepts (from applying DUCK to applying RABBIT), nor is the switch merely one in the saliency of some low-level representations (emphasizing duck-beak shape to emphasizing rabbit-nose shape). In my view, changes in low-level saliencies change the structuring process, thus changing the implicit similarity relations that will be primed for the perceived figure. If one attentionally selects the duck-beak shape, this will not only make more visually salient one particular low-level feature (that localized shape), but it will also make the schematization of the perceptual representation different. The representation will (implicitly) prime other perceptual representations that have a duck-beak shape as one of their features. In other words, the figure will then be represented as having a specific duck aspect. Then, when one switches to attentionally select the rabbit-nose shape, again not only is the structure of low-level saliencies changed, but also the whole schematization process, since now *rabbit-nose shape* is going to be the filtered feature from which similarity priming will unfold. One now sees the figure as having a rabbit aspect.

Schematization does not require the mastery of associated concepts. Suffices that one has perceived before similar objects sharing similar features to the occurrent figure and have stored their representation as perceptual memory traces. Perceptual switching can then happen automatically and involuntarily through stochastically switching attentional patterns (blind flailing), which stochastically switch similarity priming and thus aspect perception.

But crucially, we can cognitively control such schematization process, making it *voluntary* instead of stochastic. One might store a concept of DUCK which points to some observational information about what features duck usually display. Such observational knowledge might trigger a specific top-down attentional command, which modifies the schematization process in a directed way. It can be quite difficult to be sensitive to the bistability of a figure if one doesn't know beforehand between which kinds of percepts the figure can switch. Being explicitly told between which categories the ambiguous image can switch might greatly help us in making the switch happen: one goes from blind flailing to myopic flailing. If I am told that the image can also be perceived as a duck, I will be on the look for specific low-level features such as a duck-beak shape. Thus, cognition can synchronously help to narrow the search for the key switch-features to be attended to. But cognition and conceptual mastery are *not* constitutive of aspect switching. Aspect switching is rather explained by changes in structuring processes which modulate perceptual schematization. These structuring processes can unfold outside of our voluntary control.

Not only can cognition influence schematization through top-down attentional commands modifying current structuring processes, but it might also modulate the similarity priming process. Most importantly, one might *anticipate* some perceptual aspect,

thus easing the priming of specific perceptual memories when the anticipated object appears. So, in the duck-rabbit case, one might for instance believe that a DUCK is going to appear. Such anticipation might provoke some pre-activation of perceptual memories of ducks. Upon seeing the duck-rabbit figure, the similarity priming that retrieves duck memories will thus be made more probable, making in turn more probable the perception of a duck aspect. Notice that the anticipation is *not* a case of similarity priming, but one of object remembering. As long as I am seeing no picture, I cannot yet be seeing a duck aspect (though I might be implicitly recalling memories of ducks already).

Another striking illustration of conceptually-driven modification of schematization is provided by "droodle"[86] pictures as presented in **Figure 10**, below. Before reading the caption of such pictures, one most often sees them as depicting some meaningless arrangements of shapes. After having read the captions though, some kind of "aspect lightning up" is experienced much phenomenologically like for duck-rabbit pictures. Nothing changes in the picture, and yet we are seeing it differently. This lightning up of an aspect can be explained by a process whereby schematization is modulated by acquiring a new belief about the picture after having read the captions. This new belief modulates schematization in two interacting and dynamic ways: first, it might change how we attend to the picture, thereby changing the structuring process; second, it might steer similarity priming through object remembering. When one reads the first caption, one might thus both recall object memories of elephant trunks, and might attend to the image in a new way that perceptually structures the image so as to evoke their trunk, orange aspect[87]. As Linda B. Smith and Diana Heise put it succinctly:

> Explicit conceptual knowledge can push perception only so far (no amount of information will convince a person looking at a spoon that it is a tomato). However, conceptual knowledge can cause individuals to seek out perceptual dimensions that make sense of conceptual distinctions. (L. B. Smith & Heise, 1992, p. 262)

What is special about such droodle pictures is that they are too elementary to evoke any automatic schematization when first perceived. The caption serves to guide such schematization by guiding attentional commands and similarity priming. Such an explanation does not require to appeal to concept application. The aspect lightning up is a

[86] The term "droodle" (from "doodle" and "riddle") was invented by American humorist and cartoonist Roger Price, who popularized this kind of entertaining drawings in the 1950s.

[87] It is known empirically that attention can modulate represented properties in higher-level vision (Çukur et al., 2013; Nastase et al., 2017).

purely perceptual phenomenon, though the phenomenon is in this case causally triggered by the acquisition of some perceptual belief.

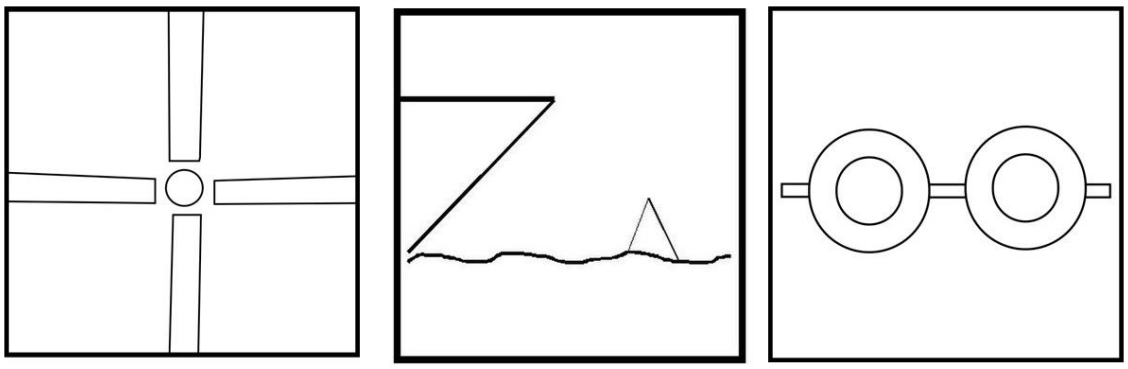

Figure 10. Three "droodle" pictures.

Now, of course, one cannot just decide to see anything as having a particular aspect through voluntary control. The particular that is perceived, and the low-level properties it instantiates, must be compatible with the attempted schematization. One cannot see the left droodle in **Figure 10** as having the aspect of two Mexicans riding a tandem bike (and likewise for all the droodles). The droodle pictures can be schematized in specific and surprising ways because they have low-level properties that are surprisingly compatible with such schematization. Schematization depends on the similarity-priming of some perceptual memories from an occurrent structured perceptual representation. Remembering or imagining something that bear no relevant similarity to what is currently perceived cannot produce a new aspect perception. Likewise, one cannot attend to relevant features of the picture that evoke specific aspects (i.e., evoke specific similarities) if the object has none of these relevant features in the first place.

An important lesson from aspect perception with droodle and bistable pictures is thus that we can voluntarily control schematization in cases where the particular to be schematized is elementary or ambiguous enough to yield some flexible kinds of

schematization (while remaining compatible with such schematization). In most cases of daily perception, perceived particulars strictly limit the types of schematizations that can be carried out. There are only a limited number of perceptual memories that are relevantly similar to a cat. One cannot normally see a cat as having a hairdryer aspect (except if some highly malfunctioning perceptual processes disrupt schematization, making one potentially misperceive a cat as having a hairdryer aspect, or misperceive one's wife as having a hat aspect (Sacks, 1998).

Controlled aspect perception with bistable figures or droodles form one of the clearest illustrations of how one can voluntarily and synchronously control one's occurring schematization processes. Such synchronic modulation of perceptual schematization can also be at work with particulars in our daily environments. This is typically the case when a novice is told where to attend to recognize a particular subordinate aspect of an object.

We can take again the example of a novice mycologist who struggles to distinguish between golden chanterelles and jack-o-lanterns. At first, the novice can only represent an undistinguished mushroom aspect. Then, after a certain period of perceptual learning guided by trial-and-error (reinforcement), the novice acquires a perceptual skill and automatizes the representation of distinct golden chanterelle aspects and jack-o-lantern aspects. Before acquiring such automated skill though, the novice might already be able to represent these distinct subordinate aspects, *but only by voluntarily and effortfully attending to specific feature dimensions*. The novice might be told by a teaching mycologist to attend to the smoother cap of jack-o-lanterns as compared to the rougher cap of chanterelles. Because we have some voluntarily control over what features we attend to (through myopic flailing), the novice will be able to follow such attentional instructions, and thus will voluntarily modify her occurrent perceptual schematization processes, emphasizing the cap-texture dimension for similarity priming. The novice might thus force a golden chanterelle aspect to be represented.

Because the synchronic top-down control of structuring through attentional modulations is coarse-grained, success in a distinctive perceptual schematization is not guaranteed. The novice might fail to attend properly to the cap-texture parts in which respect all golden chanterelles look perceptually similar. But if she succeeds, she will finally be able to perceive the mushroom as having a golden chanterelle aspect. In a way, she will have voluntarily *forced* that aspect to be represented. The novice might then have a kind of "ah-ha!" metacognitive feeling corresponding to this lightning up of a golden chanterelle aspect representation. For each new mushroom though, the novice will have to remember which attentional command ("attend to the cap-texture") she should execute to see something under a golden chanterelle aspect. If repeating in time and behaviorally successful, such controlled, top-down attentional commands might be automatized and stabilized, yielding a case of perceptual learning as we encountered in the previous section. This transition from voluntary and cognitively effortful top-down control to a fully automatized bottom-up shaping of the schematization process will allow the novice to stop having to effortfully control her attentional modulations, sparing some precious cognitive

resources that can be used elsewhere, yielding a case of perceptual offloading. She can now effortlessly and automatically form distinguished representations of golden chanterelle aspects in a fully bottom-up way, without having to think about it and to voluntarily "force" it.

There is thus an interesting gradient in how we control the schematization process. An innocent observer (one with no specific recognitional goals) blindly flails her attention around, not yielding any specific kind of perceptual schematization, or only very generic ones[88]. A training novice observer (one with specific recognitional goals but no perceptual expertise in schematization) voluntarily and effortfully ("myopically") flails her attention toward relevant feature dimensions to schematize the stimulus under relevant respects. Finally, the perceptual expert observer automatically and effortlessly structures in relevant respects for schematization. The expert *cannot help* but see some specific aspects of an object of expertise. As I said in the previous section, we are actually all experts in many kinds of schematization.

VI.2.2. Schematization Control is not Cognitive Penetration

If I am voluntarily attending to the duck-beak shape of the duck-rabbit ambiguous figure, and perceive as a consequence of this particular attentional command a duck aspect, should this be considered a kind of cognitive penetration? I have argued in II.4 Cognitive Penetration that cognitive penetration is not the kind of mechanism that could satisfactorily ground perceptual Liberalism, because there exist no uncontroversial cases of *narrow* cognitive penetration. As a reminder, a narrow definition of cognitive penetration maintains that cognitive penetration is defined by two necessary and sufficient criteria: (i) a synchronic causal effect of a cognitive representation on a perceptual representation; (ii) the causally affected perceptual representation is modified by the cognitive representation *in a semantically coherent way*. As Zenon Pylyshyn puts it:

> [I]f a system is cognitively penetrable then the function it computes is sensitive, in a semantically coherent way, to the organism's goals and beliefs, that is, it can be altered in a way that bears some logical relation to what the person knows. (Pylyshyn, 1999, p. 343)

Opposing this narrow view, there exist more relaxed *wide* views of cognitive penetration that do not require the second semantic coherence condition, and count as cognitive penetration any kind of relevant top-down alteration of perception by cognition. Dustin Stokes for instance has defended a *consequentialist* understanding of cognitive penetration (Stokes, 2015): any influence of cognition on perception that would bear important consequences for central philosophical issues (such as the architecture of the mind or the epistemology of perception) should be counted as cases of cognitive penetration.

[88] Though the innocent observer can be "attentionally lucky" and might end up attending to the same feature dimensions as the novice or the expert observer.

As I argued in II.4, only a narrow understanding of cognitive penetration could potentially vindicate perceptual Liberalism. The argument went as follows: the contents of cognitive representations such as beliefs and desires are often high-level conceptual contents. If such representations can alter perceptual representation in a semantically coherent way, then perception could inherit the high-level conceptual contents of cognition. I have argued that there exists as of today no uncontroversial empirical demonstration of such a semantically coherent effect of cognition on perception. But one might wonder whether my defense of perceptual schematization, and how it can be controlled by explicit intentions adjusting schematization, cannot be taken as a demonstration of narrow cognitive penetration.

We should note first that perceptual schematization does not necessarily depend on the top-down modulation from cognitive states. Phylogenetic selection and perceptual learning are processes that do not count as cognitive penetration on its usual understanding (either wide or narrow), because these are processes that unfold diachronically at the phylogenetic or ontogenetic scale (Arstila, 2016). I argued that perceptual learning can succeed through purely behavioral reinforcement loops (Roelfsema et al., 2010): if a schematization strategy leads to behavioral success (e.g., attending to one feature dimension to recognize preys), then this specific schematization might be stabilized without any intervention from cognition. One does not need to intend to categorize something to learn to recognize it through schematization. Such learning can be fully unsupervised.

What about cases where schematization is guided by voluntary, synchronic top-down attentional modulations as we saw in the preceding section? Shouldn't *this* count as cognitive penetration? Surely, there is a sense in which there is semantic coherence between my desire to see a DUCK and my seeing a duck aspect on the other. But as I have insisted, DUCK and duck aspect () are different representations, with different contents and a different semantics. DUCK is a concept that represents the natural kind property of *duck-hood*.  is an aspect that represents the superficial kind property of *duck-form*.

Determining whether synchronic attentional modulations amount to cognitive penetration amounts to determine whether DUCK and  are semantically coherent in a relevant way. Ultimately, I think the issue turns out to be terminological. On *very* narrow views of cognitive penetration, one can require that perception directly computes over the conceptual representations of cognition, including the amodal (verbal) information they point. For instance, E. J. Green (2020) argues that while cognition can be demonstrated to causally influence perceptual processing (through top-down attentional selections and modulations), it does not add new computational dimensions to perception, what he calls dimensional "enrichment". Also discussing aspect switching with ambiguous figures, he concludes that:

> Cognition affects which objects or features get selected for further processing, and thus affects which high-level processes dominate at a time. But there is no reason to

> think that this enriches the dimensions computable by either early vision or high-level vision. (Green, 2020a, p. 352)

For Green, the fact that perception cannot be enriched dimensionally forms a criterion for its *encapsulation*: perception can only compute over a pre-defined and limited number of representational dimensions (the realm of perceivable properties), while cognition's dimensional domain is supposedly indefinite and limitless (it is *unencapsulated*).

Green explicitly makes a distinction between debates about encapsulation on the one hand, and debates about cognitive penetration on the other, though in most traditional discussions (Fodor, 1983; Pylyshyn, 1999), these notions are taken as equivalent. One could reasonably hold a view where perception is still cognitively impenetrable even if we accept synchronic attentional modulations of its representations, if such modulations do not *add* new representational dimensions that perception couldn't have on its own. This view would require that perception directly computes over the conceptual representations of cognition for it to be penetrated. It would require perception to be able to compute over concepts like DUCK from cognitions, i.e., to use not only perceivable information, but also information accessed by semantic pointers (e.g., that 'DUCKS have a BEAK and FEATHERS'). This view of semantic coherence is thus a view that requires a literal *semantic* penetration from cognition to perception. As I have emphasized, postulating that perception can represent conceptually is a highly controversial claim, and one that we should reject if we do not want to beg the debate in favor of perceptual Liberalism. It is also knowingly hard to demonstrated such semantic effects.

One could also have a more relaxed interpretation of the semantic coherence criterion. One could be content with the idea that having a causal relation between a cognitive representation C and a perceptual representation P, *and some kind of semantic affinity between them*, is sufficient for cognitive penetration. This is exactly the position of E. J. Green: he argues that while perception is encapsulated on the dimension-restriction criterion, it can still be considered cognitively penetrated on the semantic coherence criterion. In fact, some synchronic top-down causal effects of attention on perception can be considered as cognitive penetration on this relaxed view because the cognitive input (e.g., the desire to see a DUCK) modulates perceptual representations (one's seeing a 🦆) with a certain semantic affinity: the concept of DUCK is semantically affiliated to the percept of duck aspect.

But what exactly does this semantic affinity amount to? I think the notion is condemned to remain rather vague and to rely on intuition for its tracking. It is quite intuitive that the concept DUCK and the aspect 🦆 bear some semantic affinities. Some of this affinity might have to do with the very deep causal relations that these representations entertain: acquiring the concept DUCK might rely on first being able to represent the aspect 🦆; and learning to represent some aspects. Likewise, the subordinate golden chanterelle aspect might require one to first acquire the concept GOLDEN CHANTERELLE to guide her learning process. Associations between aspects and concepts are widespread: human adults,

when perceiving a duck aspect (recognizing a duck) might instantly apply a DUCK concept to it (categorizing), and produce judgments as of a DUCK.

Furthermore, concepts understood as semantic pointers might be linked to the same perceptual memories as are primed in related aspects: my DUCK concept might retrieve the same perceptual memories as are retrieved when I schematize a duck as having a duck aspect. *Observational* concepts bear on this criterion a lot of semantic affinity with corresponding aspects. Anything that is accurately perceived as having a duck aspect might be categorized under the DUCK observational concept. Their referent (superficial kind property) is also equivalent. Nonetheless, I insist that they are not equivalent, since a duck aspect representation is a demonstrative representation that is nonconceptual (that *cannot* point to non-perceptual information stored amodally). One cannot *think* about duck aspects (one can only see them). But one *can* think about DUCKS, even understood purely observationally. One can think that 'DUCKS are CUTE' (upon seeing one for the first time, without anything non-perceivable about them). Even though 🦆 and (observational) duck might represent the same (superficial kind) property, "semantic affinity" cannot fully rely on this sameness of referent, because the (non-observational, general) concept duck might also refer to a different (natural kind) property of *duck-hood*.

For the purpose of this debate, it suffices that we *do* have the intuition that DUCK and 🦆 bear a semantic affinity, whatever this semantic affinity turns out to be. Thus, that a cognitive representation of DUCK can causally trigger our representation of 🦆 through top-down attentional modulations would fulfil the criteria of semantic coherence of this view, and thus should count as cognitive penetration.

Ultimately, whether the label "cognitive penetration" should be applied only in (hypothetical) cases of strict semantic penetration or dimensional addition (as in Gross, 2017) or should be applied for cases of mere semantic affinity (as in Green, 2020b) is a not-so-interesting terminological matter. What matters for us is to make the distinction explicit. Perceptual schematization aided by top-down cognitive modulations is *not* a case of strict semantic penetration, but it is a case of causal influence sustaining semantic affinity. Whether one wishes to call it cognitive penetration will thus depend on how strict one wants to be about the criterion of semantic coherence.

As Dustin Stokes has argued (Stokes, 2015), this definitional choice might be guided by one's explanatory goals, with the idea that accepting cognitive penetration should lead to interesting revisionary conclusions regarding the mind. In the ambit of the admissible contents of perception debate, it seems that only the *strict* understanding of semantic coherence as semantic penetration would lead to revisionary conclusions, since it would vindicate (if not trivialize) perceptual Liberalism by making perception partly a conceptual representational system akin to that of thought and reasoning. On the other hand, my argument about schematizations grounding a kind of strictly perceptual high-level representational state (aspects) is orthogonal to the idea of cognitive influence. It could have

been that perception were fully impenetrable and informationally encapsulated, that schematization would still occur, and my argument for Liberalism would still stand.

Thus, it might be more theoretically conducive in the context of this thesis about the representational architecture of the mind to keep with the very narrow understanding of cognitive penetration. As things stand, we have no reason to believe that perception is semantically penetrated by cognition (this would require to show that perception directly operates over the conceptual contents of cognition), even though it can be modulated by it.

In other explanatory contexts, the idea that cognition can causally influence perception while sustaining some kind of semantic affinity might in itself be revisionary. This might be particularly true in epistemological contexts. If one believes that perceptual representations constitute prima facie justifiers for beliefs, then the fact that such perceptions can themselves be modulated by beliefs in a semantically coherent way might pose problems of vicious justificatory circularity (Siegel, 2013b, 2017). For instance, my belief that '<u>there is a DUCK in front of me</u>' might modulate my perception so that I see a  (imagine for instance a situation where we encounter a strange animal that looks exactly like a duck-rabbit bistable figure). This latter perceptual aspect might itself justify my belief that 'there is a <u>DUCK in front of me</u>'. In this case, the etiology for the justification of my belief would be perniciously circular. Thus, in the ambit of epistemology, it might be theoretically conducive to accept the mere semantic affinity criterion for cognitive penetration. After all, one of the original reasons why the concept of "cognitive (im)penetrability" was put forth by Jerry Fodor was to preserve the epistemological reliability of perception.

VII. General Summary and Opening Themes

VII.1. General Summary

The six chapters constituting this thesis have provided what I hope to be an original and compelling defense of perceptual Liberalism. I summarize here the argumentative journey, before considering some potential consequences of my defense of Liberalism within further philosophical topics.

In Chapter I, I argue that the admissible contents of perception debate can easily be vitiated by confusions about what exactly is meant by *perception*. I try to avoid this conundrum by restricting the field of inquiry to *perceptual representations* as they are understood by contemporary practices in the perceptual sciences. I emphasize that perceptual representations should be understood as a genuine mental kind separated from *cognitive* representations. The admissible contents debate only concerns *perception-as* and is unrelated to *perception-that* (perceptual *beliefs*). We want to determine whether one can perceive some particular *as* F, where F is a high-level property. Such representational perception-as shouldn't be assimilated and reduced to perceptual experience. The relevant target of the admissible contents of perception debate is not whether we can perceptually *experience* some high-level character F*, but whether we can perceptually *represent* high-level properties. Answering this question in itself bears sufficient consequences for psychological, epistemological, and maybe ethical inquiries to deserve all our philosophical attention.

In Chapter II, I review the most prominent arguments in favor of Liberalism that have been offered in the recent philosophical literature. I argue that these arguments suffer from an overemphasis on phenomenological intuition or rely on premises that are just as (if not more) philosophically or empirically controversial as Liberalism itself. The recent batch of empirically-based "common hallmark arguments" avoid these obstacles and constitutes a clear improvement over previous argumentative strategies for defending Liberalism. Still, such common hallmark arguments can offer only *abductive* and *prima facie* support for Liberalism. A Conservative can always reasonably propose an alternative explanation of the empirical data upheld by Liberals. What explanation will be deemed most convincing seems to ultimately depend on one's prior conviction for Liberalism or Conservatism, and argumentative circularly is thus threatening the potency of common hallmark arguments. I claim that to avoid such vicious circularity, common hallmark

arguments should be complemented by an explicit demonstration of which *psychological mechanism* could explain the perceptual capacity to represent high-level contents.

In Chapter III, I set out some criteria for searching for a psychological mechanism that is uniquely *perceptual* in nature. Perceptual mechanisms are defined instrumentally in opposition to *cognitive* mechanisms. Cognitive mechanisms (typically: deliberative thought) are computational manipulations of *conceptual* representations that are subcomponents of propositional, language-of-thought-like structures with a discursive format. Oppositely, we can think of *perceptual* mechanisms as manipulations of *nonconceptual* representations that are components of a nonpropositional structures with an iconic format. This view aligns with recent philosophically authoritative and empirically-based proposals for distinguishing perception from cognition (Block, 2023; Burge, 2022). I need not commit to the idea that this criterion constitutes a necessary or sufficient condition for defining perception. I am only using it in an instrumental way: it is a sufficiently restrictive criterion so that perceptual Liberalism (and the whole admissible contents of perception debate) be not trivialized. Having this criterion in hand, I show that the empirical literature on perceptual categorization or object recognition, which investigates how "perceptual" computations allow us to classify particulars (yielding high-level representations), most often describes psychological mechanisms that violate the nonconceptuality criterion. On closer scrutiny, these models are most often "concept application" models which rely at some point on access to semantic stores of concepts which are applied to what is perceived. Different scientific models of perceptual categorization only disagree as to which mechanisms guide such access: rule-based mechanisms or similarity-based ones (or combinations of them). As such, the scientific literature at best describes *hybrid* perceptual-cognitive mechanisms of categorization, and it cannot provide us with a purely perceptual mechanism to vindicate Liberalism.

In Chapter IV, I show that a similarity-based psychological mechanism *without access to conceptual stores* remains sufficient on its own to vindicate perceptual Liberalism. I call such mechanism *schematization*. Schematization is a psychological mechanism which consists in (1) *structuring* low-level perceptual representation into a streamlined and filtered high-level representation, and (2) *priming* perceptual mnemonic traces that are most *similar* to such structured representation. Crucially, the structuring phase allows one to determine relevant *respects* for similarity, which would otherwise be indeterminate. I uphold a critical distinction between such schematization mechanisms and *categorization* mechanisms as were described in the previous chapter (including *similarity-based* categorization). While the latter are hybrid perceptual-cognitive mechanisms involving access to semantic stores, schematization is a purely perceptual mechanism. Structuring is a hierarchical bottom-up process of dimensional simplification over low-level unstructured representations within the perceptual system (though it can be modulated top-down by cognitive processes). Similarity priming only searches in *perceptual memory* for similar representations. Perceptual memory is a long-term stock of nonconceptualized mnemonic traces that can be automatically and implicitly primed by similar occurrent perceptual representations. I show

that schematization is a psychological mechanism that is strongly supported by empirical considerations: it seems necessary to account for observed perceptually-based behavioral generalizations, especially in conceptless subjects like infants and animals; it is further backed up by brain imaging and computational models (deep neural networks) of the human perceptual system.

In Chapter V, I argue that schematization grounds a specific kind of high-level representation: *aspects*. Aspects fulfil basic conditions for being high-level representations: they are vertically articulated to low-level perceptual representations (there needs to be a representational input to be schematized in the first place), but they are not reducible to them (many different low-level representations can correspond to the same aspect, and many different aspects can be articulated to the same low-level representation). I insist though that while being the representational outputs of schematization, aspects do not represent *similarity* as such. Aspects are representations of specific *kind*-properties. They represent particulars as pertaining to specific classes of particulars determined by superficial *form* similarity. Superficial similarity, and thus superficial kind properties, are objectively determined by the physical *surface body* of objects. Aspects represent such superficial kind properties. Aspects can only represent a high-level kind content that is related to the perceivable physical properties of objects. Superficial kind representations are thus to be distinguished from natural or functional kind representations. An apple aspect is not an APPLE concept. The former is veridical of plastic apples, the latter isn't. The former is a perceptual representation, the former is a cognitive representation. Perception, through schematization, deploys many representations (aspects) that are causally related to objective superficial kind properties instantiated by particulars in the environment. Two individuals can be perceptually causally related through *different* schematizations (structuring and similarity priming) to the *same* instantiated superficial kind property (e.g., the cow-form property). There are different *modes of representation* of the same aspect. I think we can be open-minded as to whether aspects should be individuated by their external representational reference (i.e., superficial kind properties), or by their internal modes of representation (i.e., by idiosyncratic types of schematizations).

In Chapter VI, I consider a potential issue for my view: how is it that perception *on its own* knows how to schematize? If schematization is purely perceptual, it cannot (by my own criterion) compute over the conceptual contents of cognition. But then how is it possible that upon encountering an apple, I automatically and effortlessly schematize it so as to perceive an apple aspect, *even if* I might not judge (or might simply be unable to judge) it is an apple? I answer this worry by showing that schematization is a perceptual mechanism subject to various prewiring and learning processes. Some schematizations might simply be prewired through classical processes of evolutionary selection. But most schematization processes are *automated* through perceptual learning: specific schematizations are learned and retained when they yield useful behaviors (including categorization) for an individual in a particular context, through reinforcement learning. While such automation of schematization can occur outside of cognitive control, this

doesn't mean that schematization is impervious to cognitive influences. On the contrary, I argue that the schematization process can be modulated in a top-down way, mostly through attentional commands or potentiation of perceptual memories. It is an interesting question to consider whether such cognitive influence on schematization should count as cognitive penetration. I think not, or at least, not on a sufficiently restricted view of cognitive penetration under which only modulations with genuine semantic (conceptual) transfers from cognition to perception should count as genuine penetration. Perception simply cannot compute over concepts (though it can be influenced by them indirectly), so no such semantic transfers are postulated by my view. Obviously, if one accepts a more relaxed criterion for cognitive penetration, then one might accept that perception is cognitively penetrable.

I now conclude this thesis by exploring some consequence of my aspect Liberalism for philosophical considerations about the mind's architecture and the epistemological role of perception.

VII.2. The Heterogeneity of Representational Types in the Mind

The mind hosts a motley collection of representational types. Aspects are not identical to concepts. Aspects are not reducible to low-level perceptual representations. One shouldn't be a representational monist about the mind. One should be a representational pluralist. An overemphasis on thought and attached linguistic-conceptual capacities obscures this pluralism. An attentive consideration of *perceptual* representational capacities puts it under a bright light. Perception doesn't have the representational capacities of cognition, and *vice versa*. A particular grasped under perceptual capacities won't yield the same representations as that same particular gasped under cognitive capacities. Four representational types have been identified in this thesis, as presented in **Figure 11** below. The central distinction that we considered was that between perceptual *low-level representations* and *high-level aspects*. Schematization is the psychological mechanism that allows one to go from the former to the latter. Seeing one and the same apple only under its low-level properties (including its specific gestalt) is not the same as seeing it under an apple aspect. John, seeing an apple for the first time, cannot schematize it under an apple aspect, since John hasn't stored any perceptual memory looking like apples. This is so even if John *attends* to the apple in the same way as Jack, who does have perceptual memories of apples and does schematize it under an apple aspect. John and Jack thus have a different perceptual representation of the apple. John only represents some low-level properties of it and maybe some generic aspect like a fruit aspect. Jack represents low-level properties of it *and* represents the high-level superficial kind property of *apple-form*.

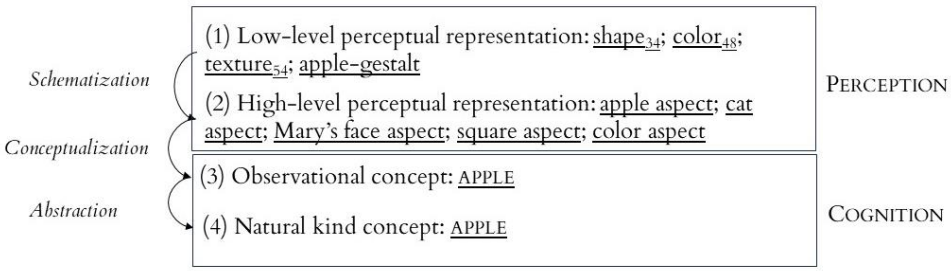

Figure 11. Four representational types in the mind.

There is thus some vertical representational heterogeneity already *within* perception. This heterogeneity continues into cognition. We come to think about what we see. At some point, one is able to betoken a specific *cognitive* representation for apples. One enters into the realm of concepts. It is intuitively probable that many concept representations in cognition are first of foremost derivative from aspect representations in perception. This seems to be how children learn their first words (expressing concepts): parents point at particulars so that their children *look*. *Observational* concepts are concepts which only carry information about, or point to, perceptual contents. Observational concepts though are not identical to their perceptual counterparts. In fact, contrary to perceptual representations, conceptual representations are constitutively subcomponents of propositional attitudes. Crucially, concepts can be manipulated in the absence of the referred to perceptual properties. Perception essentially depends on a causal relation to particulars in the environment. Thinking depends on being able to mentally break away from this online perceptual causal relation.

Conceptualization of aspects into observational concepts is a crucial step in such a breakaway. Conceptualization consists in linking perceptual information to a mental file stored in semantic memory that is itself amodal but can retrieve perceptual modal information. This file can then be non-demonstratively used in thought even in the perceptual absence of the referred to properties. Infants and animals might be able to manipulate purely observational concepts. Old enough infants and (many) animals are able to hide an apple somewhere (say at location *L*), and later on are able to think that '<u>the APPLE is at *L*</u>' so as to guide their steps back towards the apple. In this case, the <u>APPLE</u> concept can be purely observational. Infants and animals do not store more information about apples than what they look like (or taste like, smell like etc.).

Obviously, not all our concepts are purely observational. Many of our concepts (<u>PRESIDENT</u>, <u>JUSTICE</u>, <u>QUARK</u>) point to non-perceptual information. Furthermore, observational concepts can be abstracted by learning non-observational information about their referent, a process I call "abstraction" in **Figure 11**. Thus, one can go from mastering the observational concept <u>APPLE</u> to mastering the *natural kind* concept <u>APPLE</u>. This natural

kind concept preserves observational information, but adds to it further information like, e.g., EDIBILITY, NURISHING, IS LOVED BY MARY, HAS GALA AS A VARIETY etc. Once non-observational information is added, the concept APPLE can come to refer to *applehood* (the natural, fruit kind).

It is a further interesting issue to consider whether acquiring such non-observational information requires one to master a language (Gauker, 2011). Possessing purely observational concepts only requires one to have propositional thought capacities. If one follows the orthodoxy in the analytic tradition, thought capacities precede language capacities (since language *expresses* thoughts). Thus, it seems possible that a prelinguistic organism (infant and animals) could possess observational concepts. These are obviously complex matters, and I prefer to leave the question rather open here.

I thus think one should maintain a clear distinction between nonpropositional-nonconceptual perception and propositional-conceptual cognition on the other, *even though* one accepts some high-level representations in perception. This acceptance of a vertical heterogeneity of representational types stands in stark contrast with the kind of representational monism advocated by radical neo-empiricists on the one hand, and radical intellectualists on the other.

Going against neo-empiricist views (Barsalou, 1999, 2010; Prinz, 2002), nothing of what I have advocated requires that all concepts be fully grounded in perception, even if many of our concepts retrieve, *inter alia*, observational information. In the spirit of empiricism though, my aspect Liberalism asserts that perception is *on its own* capable of representing some kind of high-level contents, *by sheer association* (through schematization) *of nonconceptual representations* (or of "sensations" as a classical empiricist would have put it).

Going against intellectualist or conceptualist views (Fodor, 1998; McDowell, 1994), my view entails that there are high-level representations in the mind that are nonconceptual. *Aspects* are such representations. They are subcognitive, purely perceptual, nonconceptual representations of superficial kind properties. As we will see shortly, aspects shouldn't be an obstacle though to the idea that perception can constitute some kind of primary epistemic basis for our judgments, even if they are nonconceptual representations (Burge, 2003). Intellectualists should be reassured too.

As I have pointed out in the introduction, perceptual Liberalism has been invoked to justify the perception of a very diverse fauna of high-level contents, including natural and functional kind contents (Siegel, 2006), causal contents (Butterfill, 2009; Siegel, 2011), ensemble contents (Bayne & McClelland, 2019), aesthetic contents (Stokes, 2018b; Voltolini, 2023), moral contents (Audi, 2010; Cowan, 2015), affordances (Nanay, 2011), emotional expressions (Newen, 2017), intentions (Pacherie, 2005) or even personality traits (Brogaard, 2016).

An advantage of my defense of perceptual Liberalism through an expounding of schematization mechanisms is that it provides us with a general criterion to determine which high-level contents are or aren't representable in perception. Schematization necessarily works with low-level perceptual inputs to prime similar perceptual memory

traces. Resulting aspects can only have *superficial kind properties* as their content. For each proposed high-level representation in perception, one should thus first ask whether it is possible that one can schematize it. A good rule of thumb here is to wonder whether all particulars to which we attribute a superficial kind property have some distinguishable surface body similarity. If not, they cannot enter into schematization processes. As we saw, this criterion allows to rule out natural kind and functional kind properties as potential candidates for being perceptible through schematization. In fact, there is nothing distinguishable in terms of the surface body of particulars belonging to a natural kind as compared to lookalike particulars not belonging to this natural kind (e.g., apples vs. plastic apples). However, we can perceive *superficial kind property* counterparts of natural and functional kind properties, such as apple-form, pine-tree-Eform or hairdryer-form.

This criterion reasonably allows us to rule out other proposed properties for being perceptible. Moral properties do not form a superficial kind and are thus not schematizable. *Thin* aesthetic properties like beauty, sublimity or prettiness are also probably too internally variegated to form superficial kinds. Thicker aesthetic properties like "being an Impressionist painting" are more plausibly schematizable (Stokes, 2018b). It is plausible to maintain that all Impressionist paintings are surface-alike in some respects. One might thus come to learn to schematize Impressionist paintings and perceptually represent Impressionist painting aspects[89]. This might be particularly relevant for an art expert who needs to discriminate Impressionist paintings on a daily basis. What property is represented by an Impressionist painting aspect is of course *not* the historical or artifactual property of having been painted by an Impressionist painter in France in the last quarter of the 19th century. This bundle of properties obviously does not constitute a superficial kind. A well-made forgery of an Impressionist painting made in the 21st century by a robot in Singapore *does* instantiate the superficial kind property of *Impressionist form*. An Impressionist painting aspect is thus accurate of it.

The general point of course applies to all other potential perceivable high-level properties: one can only perceive the superficial kind property of a particular, not any deeper under-the-skin essence. Thus, one cannot perceive *causation* per se (understood as some physics phenomenon), but causal spatiotemporal forms (such as *pushing-form*, *pulling-form* etc.). The same lesson should be drawn for properties such as relational properties ("*insideness-form*"), ensemble properties ("*density-form*"), affordances ("*graspable-form*"), emotions ("*anger-form*"), or maybe personality traits ("*agreeable-form*").

Of course, showing that a superficial kind property is *potentially* representable in perception does nothing to show that it actually is represented by any particular individual. What is further required is some actual schematization capacity, or schematization skill.

[89] This idea that we can perceptually represent some artistic styles might be of great consequence for some theories of aesthetic experiences. This is especially true of Kendall Walton's famous view that how we perceive a work of art depends on the category we see it as belonging to (Ransom, 2020c; Walton, 1970). I think such "categorization" can be carried out by schematization.

This skill depends both on the ability to structure the representation in relevant ways and on the storing of relevant perceptual memories. Some schematizations require a good amount of perceptual learning and constitute a kind of expert mental skill. No two individuals have the exact same schematization capacities. What schematizations are performed, and what aspects are represented is ultimately an empirical issue, solved by considering the idiosyncratic perceptual capacities of an individual.

Accepting aspect representations in perception should make us reassess the role of perception in the mind's economy. One cannot take perception merely as a passive input system to the rest of the mind. Perception computes over sensory inputs in ways that allow us to represent superficial kind properties of the environment, which properties are cognitively and behaviorally important to us. This ability does not appeal to anything like propositional attitude inferential transitions or concept manipulations. This ability is purely perceptual. From aspect representations, one can already "know" many things about the world, and one can act in quite complex ways. When Al sees Oli, Al runs towards him in joy, because Oli is Al's brother. Al does this *without thinking about it*: Al has perceptually represented an Oli aspect, and this aspect triggers a cascade of affective (joy) and action events (running). This series of states and events can be triggered without Al having to think propositionally, in his own language-of-thought, that 'this is OLI'. Al *just has to look and see*. Al could be a 2-years old child or an ape for that matter. Al could still run towards Oli because he sees someone with an Oli aspect.

Thus, aspect Liberalism refocuses our attention on the centrality of perception in explaining human behavior[90], including *cognitive* (and *epistemic*) behaviors. Many of our daily actions might flow directly from what we are seeing, not from what we are thinking. This idea is made plausible by aspect Liberalism, since it would be doubtful that representing merely low-level properties would allow us to act in very complex ways. But if we can represent high-level *superficial kind* properties, the repertoire of rational actions that we can perform is largely expanded. Just by looking and seeing, one can represent aspects of loved ones, dangerous foes, delicious foods, useful objects, or safe shelters, and act accordingly. Because schematization is a *perceptual* capacity, seeing these aspects and acting according to them is *fast*, *frugal* and *reliable*. Thinking is a recent luxury at the evolutionary timescale. It is slow, cognitively costly, and prone to many biases and errors. It can only be allowed by species who have built for themselves an environment that ensures them time and safety to do so. For most of the minded animal kingdom, it is first and foremost *perception* that guides behavior, not thinking. This does not mean that thoughtless animals are dumb, mindless automatons. Perception *is smart*. Through schematization, it carves the world at relevant superficial kind joints, allowing perceivers to act cleverly. Because schematization processes are prewired through evolutionary

[90] This idea is much in the spirit of (Nanay, 2013), though I do not, like him, focus on *pragmatic* representations of "action properties". Aspect Liberalism though is fully compatible with the idea that we can perceive *aspects* as of (superficial kind) action properties, or aspects of (superficial kind) affordances.

processes or learned through lifetime reinforcement processes, we can expect that most aspects will represent behaviorally useful superficial kind information for the perceiver.

Reliance on aspect perception is not limited to thoughtless organisms. As highly advanced thinking animals as 21st century *Homo sapiens* also schematize a whole lot. They do so without having to think about it. Most schematization is automatic (except those schematizations that are top-down guided by cognition, especially when stimuli are impoverished or ambiguous). Again, this is rational considered from the point of view of the economy of cognitive resources. Many of our actions only require that we represent some superficial kind properties. To grasp the fork in front of me, it is sufficient for me to perceptually represent a fork aspect, and to have associated with such aspect some specific action routines. It would be costly to have to *think* 'this is a FORK' each time one wants to grasp a fork. I can grasp a fork in the cutlery drawer without having to think about it, and that is a good thing. While I grasp the fork, I can think about what present I should get for Mary's birthday. The task of recognizing a fork has been offloaded to perception (to schematization). Representing a fork aspect is here sufficient (at the perceptual level) for triggering the simple action of grasping (in addition to maybe representing the fork's affordances). Tokening the concept FORK (i.e., thinking) is irrelevant and redundant, and is a waste of cognitive resources. Of course, at a meeting of cutlery businessmen, thoughts of FORK are going to be prevalent: are FORK sales going to raise next year? What does the customer look for in a FORK? Is there a place for luxury FORK in the business? Etc. Linguistic communication is a quite complex behavior, and one that requires the mastery of concepts. In this situation, aspect representations are insufficient, and propositional-conceptual processes are required.

But in many behavioral situations, aspect perception is enough to act, and so we better transfer cognitive load to fast, frugal and reliable perceptual schematization. We better *just have to look* to know what to do. This is something chess players, radiologists, ornithologists, mycologists, gemologists, car dealers, and many other perceptual experts (implicitly) know. They have offloaded a cumbersome cognitive process (e.g., knowing that 'the Scotch Opening is the chess opening that begins with the moves 1. e4 e5 2. Nf3 Nc6 3. d4') to perception (e.g., one has learned to directly *see* a Scotch Opening aspect). We are *all* perceptual experts of some kind. We can all expertly schematize for perceptually representing the specific face-aspects of family members, colleagues, and friends. We are all perceptual experts for the schematization of aspects of daily objects: clothes, road signs, foods, cutlery, appliances, electronics, books, buildings, and whatnot. This field of aspects allows us to behave at a cheap cognitive cost. *One just has to look*.

Aspect representations are ubiquitous *tout court* since schematization is automatic. *How* we schematize might be highly idiosyncratic and subject to top-down modulations, but one cannot impede schematizing. Schematization is in this sense not just automatic but *ballistic* (Mandelbaum, 2015). Perception-without-schematization, and thus aspectless perception, is at the very least extremely rare, if not totally impossible (for human adults at least). We seem to ubiquitously see things under specific aspects (beside their low-level

properties), even when we are not performing any action (this is obvious for instance, when one is watching a movie, or passively contemplating a landscape).

I do not wish to claim that aspect representations are essential merely for performing simple actions such as grasping, eating, searching for mushrooms, or running towards Oli. Aspects might underlie some much more complex endeavors, such as going to an art exhibition or looking to buy a new couch for the living room. Again, this is about fighting a battle against overintellectualization. Obviously, some propositional thought is going to be involved in going to an art exhibition (e.g., thinking about the intentions of the artist) or looking for a new couch to buy (e.g., thinking that leather is not a comfortable fabric and should be avoided). But it is not *all* that is going on. One also goes to the museum or to the furniture shop to *look*. Aspects are sought for themselves. One might be an admirer of Impressionist paintings, and more precisely be an admirer of Impressionist painting aspects. One might be such an admirer without knowing anything about Impressionism as a historical movement. One might be looking for a very specific superficial kind property for her living room couch (a form she has learned to schematize and thus recognize by hours of searches in furniture magazines). Aspect perception is thus not only important for simple, "thoughtless" behaviors we have in common with many animals, but also for behaviors we regard as paradigmatic of a sophisticated human life, such as going to an art exhibition or choosing a design fitting couch for our living room. We find some aspects pleasing and are seeking them. We find others unpleasing and try to avoid them. While aspects are perceptual in nature, they stand in tight connections with all sorts of thinking, feeling, and acting, from the most humdrum human behaviors to the most dignified human enterprises.

VII.3. Schematization and Epistemic Warrant

Among those highest human endeavors where aspects play a crucial role, that of *truth-seeking* deserves our closing considerations. As we saw in the Introduction, one of the main reasons why the admissible contents of perception debate drives so much philosophical interest is that many philosophical theories take it that perception plays a crucial epistemic role in warranting our beliefs, and thus ultimately plays a crucial role in grounding knowledge. By extending perception's representational paraphernalia to *aspects*, the defended Liberalism here is expected to also extend the number of our prima facie warranted beliefs. If I see an apple aspect in front of me, it seems I am at least warranted in holding the basic observational belief that 'something looking like an APPLE is in front of me'. In fact, the information that is pointed by my APPLE observational concept matches the superficial kind property represented by the apple aspect representation. Of course, this warrant does not require any kind of perceptual infallibility. One might just be having some kind of perceptual illusion. But it suffices that perception be reliable in some normal circumstances to count it as providing some epistemic warrant for our observational beliefs.

Furthermore, aspects might come to warrant not just purely observational beliefs, but also the most classical kind of general beliefs involving general concepts, such as that 'an APPLE is in front me', here in the sense that one believes that the natural kind property *apple-hood* is instantiated. Such a transition depends on some kind of normality conditions too: one *wouldn't* be prima facie warranted to believe so if one were (unbeknownst to her) living in an environment where things that look like apples are actually all plastic apples. One can know that defeaters are present, and thus refrain from drawing a particular belief from what one is seeing: if one sees an apple aspect in a toy store, one can inhibit the transition to believing that *apple-hood* is instantiated. But notice that one is still warranted in believing that 'something looking like an APPLE is there', in the observational sense.

Transitions from aspect perception to non-observational beliefs can be epistemically warranted because of the general metaphysical principle that *how things look usually corresponds to what things are*. Indistinguishable reproductions like plastic apples or holograms which break this rule are unnatural and uncommon. The representational translation[91] between seeing an apple aspect to believing that 'there is an APPLE' (in the natural kind sense), is thus epistemically warranted in our common natural environment. This is of course a *defeasible* warrant, but it is probably the most a psychologized epistemology can get us (Burge, 2003; Goldman, 1976).

I here follow Burge (2003) in distinguishing between the kind of epistemic warrant provided by perception – what Burge calls "perceptual entitlement" – from the kind of epistemic warrant provided by *reasoning* – what Burge calls "justification". The kind of warrant provided by perception is an *externalist* warrant: perception warrants our belief because perception is a reliable representational mechanism. This is so even if the perceiver does not have the conceptual capacities to articulate the warrant (Burge, 2010). As Burge puts it:

> Entitlement is epistemically externalist inasmuch as it is warrant that need not be fully conceptually accessible, even on reflection, to the warranted individual. The individual need not have the concepts necessary to think the propositional content that formulates the warrant. (Burge, 2003, p. 504)

Oppositely, *justification* is epistemically *internalist*, since it is "warrant by reason that is conceptually accessible on reflection to the warranted individual" (*ibid*., p. 505). Thus, one is *entitled* in believing that 'there is an APPLE' from seeing a 🍎, but one is not *justified* in believing that 'there is an APPLE' from seeing a 🍎. One would be justified in believing there is an apple only if one were in addition able to provide relevant reasons for her belief (e.g., she could reason that she is seeing something that looks like an apple, that she remembers having bought this apple from the supermarket yesterday, that she has no good

[91] This transition of course need not correspond to any *conscious* transition in our mental life. It is dubious that we have some temporally organized experience of such transition. Rather, it seems that we consciously experience these mental states in a concurrent way: we perceive an apple aspect and believe that 'there is an APPLE' *at the same time*.

reasons to think the supermarket sells plastic apples etc.). Just like any other perceptual representations, aspects can thus only provide one with *externally warranted* or *entitled* beliefs. Whether this is enough for constituting a genuine kind of *knowledge* is a controversial issue, and one that I will not consider here. Still, we can say that aspects are epistemically virtuous since they at least provide us with some basic epistemic warrant that trades on the (defeasible) representational reliability of perception (just like low-level representations).

Since we are dealing with an epistemic warrant that trades on some kind of *reliability*, such warrant can come in degrees. Thus, the warrant that aspects provide for observational beliefs will be the strongest, probably as strong as any epistemic warrant can get. This is because observational beliefs (involving the deployment of observational concepts) stem directly from perception. The epistemic reliability of observational beliefs is thus *almost* equal to the epistemic reliability of the perceptual state itself, that is, very high (or at least, as high as a humans' representation can get). The reliability of the observational belief is not exactly *equal* to the reliability of the perceptual state because it might be that one perceives accurately but does not draw the good observational beliefs from perception (does not retrieve the good observational concepts). It is conceivable that one accurately sees something as green but inaccurately thinks that 'it looks RED'. Likewise, it might be that one accurately *sees* someone as having a Joe Biden aspect, but inaccurately thinks that 'this person looks like DONALD TRUMP'. Cases of associative agnosia might be cases of such semantic and epistemic failure. Observational beliefs are thus slightly less reliable than aspect representations because there is one more stage (in addition to perception) at which they can epistemically fail: namely, the transition from perception to cognition (what I have referred to as *conceptualization* in **Figure 11**). Still, such observational beliefs are probably the most reliable, and thus the epistemically most warranted beliefs we can ever have considering our fallible and finite mental capacities.

In the same spirit, *abstract*, *non-observational* beliefs are slightly less reliable than purely observational beliefs or aspect representations. In fact, they introduce yet other potentialities for epistemic failure. One might accurately see someone as having a Joe Biden aspect, accurately think that 'that person looks like JOE BIDEN', but inaccurately think that 'this is not JOE BIDEN'. One might just wrongly think that this Joe Biden looking person is a good impersonator (while actually it is really Joe Biden). Conversely, one might accurately see someone as having a Joe Biden aspect, accurately think that 'that person looks like JOE BIDEN', but inaccurately think that 'this is JOE BIDEN'. After all, this *might* be a good Joe Biden impersonator!

Non-observational beliefs (e.g., those about personal identity) can go wrong in more ways than observational beliefs can because their application depends on some further background beliefs that can themselves go wrong. But, to repeat the metaphysical mantra, *because how things look usually correspond to what things are*, transitions from aspects to observational to non-observational "abstract" beliefs should usually be epistemically reliable (we don't after all live in a world full of impersonators, papier-mâché reproductions and

holograms). It thus seems reasonable to claim that even general beliefs with kind concepts like JOE BIDEN, APPLE or CAT as their content can often be epistemically warranted by corresponding aspect representations.

Thus, I think that aspect Liberalism does extend a great deal the paraphernalia of our epistemically warranted beliefs, or at least, of those beliefs that we are *entitled* to hold from what we see (even if this is insufficient to say we are *justified* in having them). The set of those entitled beliefs is not limited to purely observational beliefs. It extends also to non-observational beliefs that are grounded in perceptual capacities, such as beliefs about CATS or HAIRDRYERS. Contrarily, highly abstract beliefs about the best form of government are much less epistemically warranted from what one perceives. I thus think one can be epistemically warranted in believing that 'the CAT is on the MAT', that a painting 'is IMPRESSIONIST', or that 'the FENCE is JUMPABLE', *just by looking*. This epistemic warranting power is made stronger by the fact that perception can represent aspects, which directly warrant observational beliefs about high-level superficial kinds, and also indirectly warrants non-observational beliefs about natural or functional kinds.

Actually, one might hold another epistemological story about perceptual warrant while holding unto perceptual Conservatism. In fact, one might argue that one is warranted in holding beliefs about the cat on the mat because one sees some *low-level* properties that one believes are diagnostic of a cat on the mat. I think that this kind of epistemic warrant would be weaker than the one allowed by aspects. In fact, the point is that perception is more epistemically reliable than cognition: perception (and schematization) is shaped by learning processes that are subject to reinforcement mechanisms that are generally reliable, whereas cognition is subject to many documented biases. Believing that there is a cat on the mat from seeing a cat aspect is more epistemically reliable than believing that there is a cat on the mat from seeing some low-level properties, because in the latter case one will have to store some complex inferential rule for interpreting highly variable low-level properties, while in the former case one just has to match a stable cat aspect with a stable (observational) CAT concept (both represent the same thing). In other words, the transition from perception to cognition (conceptualization) will be more indirect and cognitively complex on a Conservative view of perception and will have more chances to go epistemically wrong. Oppositely, the transition from aspects to concepts under a Liberal view might be viewed as noninferential: one associates some concept with some aspect without relying on background knowledge. The aspect might just be the kind of information that the observational concept retrieves (though I do *not* think that deploying the concept CAT requires simulating the cat aspect, as neo-empiricists do).

While perceptual Liberalism gets us an expansion of the basis of prima facie warranted beliefs, there might be some devils hidden in the epistemological details. The several names under which the devil is known are that of *relativism, constructivism* or even *solipsism*. If aspect representations constitute good epistemic warrants for many of our beliefs, a philosophical tension arises from the fact that many of our aspect representations are *idiosyncratic*. We do not perceive the same aspects, because most aspects are learned

during one's lifetime. Since humans live in highly diversified environments (both across geography and history), it should be expected that humans will learn differentiated schematization processes, subserving various behavioral goals that are relevant in their environments. Surely, seeing highly specific berry aspects is more relevant for the Paleolithic gatherer than seeing computer aspects (and oppositely for a Ph.D. student in 2024). And even within one society at one time, individuals still have highly diversified lives: we should expect that the Impressionist art lover and the sports car enthusiasts won't learn the same schematization processes. We should thus expect highly pluralistic, idiosyncratic aspect representations across individuals, *even in front of the same objects.*

The conclusion to be drawn here is that each and every one of us is warranted in believing different things. Epistemological considerations do not stand comfortably within such a pluralistic situation. Truth is One, and epistemic warrant should be sharable. If it is not, science couldn't proceed by sharing observations. *Everyone would have their own idiosyncratic observations, their own idiosyncratically warranted beliefs, their own idiosyncratic theories.* This issue strongly preoccupied Paul Churchland and Jerry Fodor at one point in the 1980s (Churchland, 1979, 1988; Fodor, 1983, 1984). Churchland thought that perception was representationally liberal and was constantly penetrated by our beliefs and theories, so that perception was irremediably idiosyncratic. He concluded from this thought that knowledge was bound to be relativized to specific (groups of) observers, which led him to embrace some kind of epistemological constructivism and relativism. On the other hand, Fodor thought that perception was representationally rather conservative (at least, early perception) and that it was cognitively impenetrable. He concluded from this that perceptual representations should be universally shared (at least for subjects with normal perception), and thus that we could maintain some kind of epistemological objectivism and universalism.

Aspect Liberalism offers a theory of perception that embraces idiosyncratic representations while being able to maintain some kind of objectivist epistemology. Churchland is right in insisting that perception is idiosyncratic, and that it is influenced by what we believe (though I am more prudent than him in counting any top-down cognitive influence on perception as cognitive penetration). He is wrong though in thinking that perception is so idiosyncratic and porous as to be penetrable by such abstract, heavily conceptualized beliefs as scientific theories. It does not make sense to claim that physicists "do not observe the western sky redden as the Sun sets. They observe the wavelength distribution of incoming solar radiation shift towards the longer wavelengths" (Churchland, 1979, p. 30), or at least, it does only if one takes "observe" as meaning something like "judging from what one sees", i.e., *seeing-that*. This is a confused way of talking about the idiosyncrasy of perception. Saying that what we *judge* or *believe* from what we see is idiosyncratic and porous to our background beliefs is a truism. What *is* non-trivially idiosyncratic and porous in perception is that we can learn to perceive and schematize in different ways, through diachronic perceptual learning or synchronic cognitive modulations. What *aspects* we see will thus be highly idiosyncratic and porous.

But this kind of psychological idiosyncrasy and porousness does not give credit to Churchland's epistemological constructivism and relativism. Such epistemological conclusion only ensues if our representations of the world were incommensurable. Churchland thought that the theories different individuals uphold could be totally incommensurable, and likewise perceptual states penetrated by such theories. While one could accept this broadly Quinean-Kuhnian view of beliefs and scientific theories being incommensurable because their meaning is irremediably inscribed within an idiosyncratic network of background beliefs and theories, this does nothing to show that *perceptual* representations are likewise incommensurable. At least, Churchland has not shown this. If what is highly idiosyncratic and porous in perception is *schematization* and thus *aspect representations*, then we have no reason to think that perceptual representations are in any way incommensurable. We can uphold the idiosyncrasy and porousness of perception while still accepting some kind of objectivist epistemology. How is it so? It is so because *aspects are commensurable*. Unlike (maybe) beliefs and scientific theories, aspects, being perceptual representations, are causally locked to properties of the world (in this case: superficial kind properties). Properties of the world cannot be contradictory by any metaphysical stance I know of. Thus, two perceivers might have very different accurate aspect representations of the same particulars, *but this does not mean that these representations are incommensurable or incompatible*. It's just that these two perceivers slice up the world differently: they schematize differently and are thus perceptually sensitive to different superficial kind properties. John sees a fruit aspect, Mary sees an apple aspect, and Mookie sees no particular aspect, but only low-level properties (Mookie is a baby monkey that hasn't learned to schematize apples in particular ways yet). *These are perfectly commensurable representations*. Some two scientists might uphold incompatible theories, which potentially make them schematize in different ways the same observed data (because maybe each theory is bound to particular perceptual habits, such as paying attention to different dimensions of the stimulus). But again, this doesn't mean that the aspects that are represented would be incommensurable. Ultimately, scientists can potentially learn the schematization skill to perceive the aspect in the data she is not yet perceiving. Scientific training should allow one to perceive the relevant aspects of one's observed data. It would be absurd for a gastroenterologist to tell the radiologist that because he sees no tumor aspect on the x-ray image, then there is no tumor. It's just that he hasn't learned the perceptual skill to see it.

The practice of science relies in part on being attuned to relevant superficial kind properties, i.e., membership within groups of superficial similarity (Quine, 1969). One expects that superficial kind properties should correspond to some *natural* kind property. We were first historically attuned to the natural kind *gold* by perceiving some yellow, solid, shiny contents, which were the dimensions for schematizing "gold" aspects associated with the GOLD concept. Then, scientific progress in mineralogy taught us that such aspects did not correlate well with the natural kind property of *gold-hood*. Instead, gold aspects from *microscope* observations were to be used. This does not mean that perceptual representations

of a 21st century mineralogist are incommensurable with the perceptual representations of a Babylonian mineralogist, even though their theories of what gold is might be. After all, the ancient Babylonian could potentially learn to see the same gold aspect as the modern mineralogist if he were given a microscope and some training. And the modern mineralogist still relies on the same "macroscopic gold" aspect (instead of the "microscopic gold" aspect) as the ancient Babylonian in his daily life (e.g., when she shops for earrings).

These different aspect representations simply pick out different superficial kind properties in the structure of the *same* world. So, to repeat, accepting the kind of perceptual idiosyncrasy or pluralism that aspect Liberalism implies should not lead one to embrace the kind of radical epistemological relativism and constructivism that Churchland advocates for.

The point generalizes beyond reflections about the epistemology of scientific observations. Perceptual idiosyncrasy is widespread, but it's only because we all pick out different superficial kind properties of the world, depending on our own particular interests. A forester, a radiologist, a chef, a free jazz enthusiast, a mushroom connoisseur will all represent different aspects. Schematization has been selected in light of allowing for efficient, rewarding perception-based behavior. What "efficient, rewarding behavior" amounts to is a highly idiosyncratic matter. We all have differentiated goals, reflected by differentiated perceptual representations. Again, none of this suggests that we should embrace a relativist and solipsistic epistemology. Veridicality is not relative, *but what aspects we are perceptually capable of representing is.*

Anyway, I think perceptual idiosyncrasy should not be overemphasized. Aspects are learned, and learning is context-dependent. But, if we at least fixate on a specific time in a specific society, the learning context will be vastly commonly shared. We should thus expect that we will share a great number of aspect representations. This is obviously true for particulars that populate our daily lives and with which we are all constantly interacting: cars, computers, cellphones, trees, cats, dogs, coffee cups, pens, etc. etc. (though the *mode of representation* of such aspects will greatly vary, since we have highly inter-individually varied ways of structuring and highly varied memory traces for such particulars). In many daily situations then, and across our lifetime, we can expect that we will perceive the same aspects. If we never did so, our powers of communication would be greatly decreased.

Because perceptual idiosyncrasy does not imply epistemic relativism or constructivism, this means aspect representations can be judged by norms of veridicality and rationality. Aspects can go wrong. If one sees something as having a gun aspect in someone's hand while it's simply a cellphone, one has violated the veridicality norm for aspect representations. A cellphone does not have the (objective) superficial kind property of *gun-form* (at least relative to some normal standard). Let's imagine that one has this wrong aspect representation because one has a racist (implicit) bias that associates Black individuals with guns. This association might make this person disproportionately attend more to feature dimensions that are typical of guns when seeing a Black person, thus

making him schematize the cellphone so that he sees a gun aspect instead of a cellphone aspect.

On a Fodorish[92] view, this person would not be morally accountable for her representation, because it is a perceptual representation that is encapsulated and thus not subjectable to modification by the individual, even if she wishes to. One cannot be blamed for what one cannot change (like one cannot be blamed for falling to perceptual illusions). On a Churchlandish[93] view, this person might be morally accountable, but it would be hard to say that this person commits any kind of representational mistake, since in this person's own (racist) representational economy, it is perfectly coherent that when structured in such a way, the cellphone has a gun aspect.

I think both views are problematic and run against our intuition about the case. Aspect Liberalism accounts both for the subjection of such representation to norms of veridicality *and* for the moral accountability of the perceiver for it. Aspects represent objective superficial kind properties. Representing a gun aspect when actually perceiving a cellphone is committing a mistake (relative to some standard). But aspects are highly malleable: schematization can come under processes of perceptual learning. Aspect representations thus obey norms of rationality and attached epistemic duties like other malleable representations. One can be rationally demanded not to acquire false beliefs (for instance, by consulting biased informational sources). Likewise, one can be rationally demanded not to perceptually learn to represent inaccurate aspects according to some standard norms within a society. The case of someone who unduly attends to a specific feature dimension while unduly disregarding others, prejudicing his perceptual schematization process, is a case where one violates such rationality norm and epistemic duties. In fact, the schematization learning process, which here depends on having acquired a prejudiced association between Black person and guns because of racist background beliefs, violates basic principles of rationality, such as perceptual learning being required to be guided by an unbiased data source (supposedly, one learned racist associations because one learned them from some racially biased data sources). Just like one is morally, rationally, and epistemically accountable for his false racist beliefs, one is also morally, rationally, and epistemically accountable for her inaccurate racist aspect perceptions. Aspect Liberalism dovetails with our intuitions on this matter.

Aspects are probably the most malleable representations of perception. This echoes a point made by Fodor, who distinguished between the modularity of early perception ("input systems") from the non-modularity of high-level perception. Contemporary experimental research shows that Fodor was wrong: perception in general (including "input systems" of early perception) are subject to perceptual learning throughout one's lifetime. But we can keep something of Fodor's idea: that higher-level vision is *much more*

[92] I am not claiming that this is how Fodor himself would have interpreted the case. This is just a useful idealization here.

[93] *Ibid.*

malleable than lower-level perception. Changing one's perceptual representations of distance, colors, sizes, motions, and the like is difficult and unlikely in ecological situations (though it is ultimately possible, as demonstrated by laboratory experiments). In many situations, learning to see new aspects only requires to perceptually structure the input differently and to store new perceptual memories. So, there might be a Fodorish point to make here: we are somewhat *more* accountable for our high-level aspect representations than for our low-level representations.

Aspects are the first form of high-level representations in the mind. They are what put us – animals, infants, and adults alike – in contact with behaviorally crucial properties of the world: superficial kind properties. Aspects are fundamental in achieving many of our highest intellectual endeavors: the acquisition of concepts, the warranting of beliefs, the building of scientific knowledge, the springing of aesthetic taste. They form a missing link between low-level perception and cognition. They are the first form of "abstraction" in the mind, though they remain non-conceptual, non-propositional, representations. Aspects remain purely perceptual. They are the highest contents eyes can see.

BIBLIOGRAPHY

Adams, F., & Aizawa, K. (2021). Causal Theories of Mental Content. In E. N. Zalta (Éd.), *The Stanford Encyclopedia of Philosophy* (Fall 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2021/entries/content-causal/

Arstila, V. (2016). Perceptual Learning Explains Two Candidates for Cognitive Penetration. *Erkenntnis*, *81*(6), 1151-1172. https://doi.org/10.1007/s10670-015-9785-3

Audi, R. (2010). Moral Perception and Moral Knowledge. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *84*, 79-97.

Ayzenberg, V., & Behrmann, M. (2022). Does the brain's ventral visual pathway compute object shape? *Trends in Cognitive Sciences*, *26*(12), 1119-1132. https://doi.org/10.1016/j.tics.2022.09.019

Barron, H. C., Vogels, T. P., Behrens, T. E., & Ramaswami, M. (2017). Inhibitory engrams in perception and memory. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(26), 6666-6674. https://doi.org/10.1073/pnas.1701812114

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577-660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W. (2010). Grounded Cognition : Past, Present, and Future. *Topics in Cognitive Science*, *2*(4), 716-724. https://doi.org/10.1111/j.1756-8765.2010.01115.x

Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2020). Capturing human categorization of natural images by combining deep networks and cognitive models. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18946-z

Bayne, T. (2009). Perception and the Reach of Phenomenal Content. *The Philosophical Quarterly (1950-)*, *59*(236), 385-404.

Bayne, T. (2016). Gist! *Proceedings of the Aristotelian Society*, *116*(2), 107-126. https://doi.org/10.1093/arisoc/aow006

Bayne, T., & McClelland, T. (2019). Ensemble representation and the contents of visual experience. *Philosophical Studies*, *176*(3), 733-753. https://doi.org/10.1007/s11098-018-1037-3

Baz, A. (2020). *The Significance of Aspect Perception : Bringing the Phenomenal World into View* (1st ed. 2020 édition). Springer Nature Switzerland AG.

Beck, J. (2018). Marking the Perception–Cognition Boundary : The Criterion of Stimulus-Dependence. *Australasian Journal of Philosophy*, *96*(2), 319-334. https://doi.org/10.1080/00048402.2017.1329329

Beck, J. (2019). Perception is Analog : The Argument from Weber's Law: *The Journal of Philosophy*, *116*(6), 319-349. https://doi.org/10.5840/jphil2019116621

Berger, J., Nanay, B., & Quilty-Dunn, J. (2018). Unconscious Perceptual Justification. *Inquiry: An Interdisciplinary Journal of Philosophy*, *61*(5-6), 569-589. https://doi.org/10.1080/0020174X.2018.1432413

Bergson, H. (2010). *Matter and memory* (N. M. Paul & W. S. Palmer, Trad.). Digireads.com.

Biederman, I. (1987). Recognition-by-components : A theory of human image understanding. *Psychological Review*, *94*(2), 115-147. https://doi.org/10.1037/0033-295X.94.2.115

Block, N. (2014). Seeing-As in the Light of Vision Science. *Philosophy and Phenomenological Research*, *89*(1), 560-572. https://doi.org/10.1111/phpr.12135

Block, N. (2023). *The Border Between Seeing and Thinking*. Oxford University Press.

Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2016). Concepts as Semantic Pointers : A Framework and Computational Model. *Cognitive Science*, *40*(5), 1128-1162. https://doi.org/10.1111/cogs.12265

Boyd, R. (1991). Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *61*(1/2), 127-148.

Bracci, S., & Op De Beeck, H. P. (2023). Understanding Human Object Vision : A Picture Is Worth a Thousand Representations. *Annual Review of Psychology*, *74*(1), 113-135. https://doi.org/10.1146/annurev-psych-032720-041031

Bracci, S., Ritchie, J. B., & de Beeck, H. O. (2017). On the partnership between neural representations of object categories and visual features in the ventral visual pathway. *Neuropsychologia*, *105*, 153-164. https://doi.org/10.1016/j.neuropsychologia.2017.06.010

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity : Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4. https://doi.org/10.1167/11.5.4

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325-14329. https://doi.org/10.1073/pnas.0803390105

Brogaard, B. (2016). Perceptual Appearances of Personality. *Philosophical Topics*, *44*(2), 83-104.

Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *The Journal of Abnormal and Social Psychology*, *42*(1), 33-44. https://doi.org/10.1037/h0058484

Buras, T. (2011). Reidian Dual Component Theory defended. *Canadian Journal of Philosophy*, *41*(sup1), 4-24. https://doi.org/10.1080/00455091.2014.897478

Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*, *4*(1), 73-122. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Burge, T. (1986). Individualism and Psychology. *The Philosophical Review*, *95*(1), 3. https://doi.org/10.2307/2185131

Burge, T. (2003). Perceptual Entitlement*. *Philosophy and Phenomenological Research*, *67*(3), 503-548. https://doi.org/10.1111/j.1933-1592.2003.tb00307.x

Burge, T. (2009). Five Theses on de Re States and Attitudes. In J. Almog & P. Leonardi (Éds.), *The Philosophy of David Kaplan* (p. 246-324). Oxford University Press.

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

Burge, T. (2014). Reply to Block : Adaptation and the Upper Border of Perception. *Philosophy and Phenomenological Research*, *89*(3), 573-583. https://doi.org/10.1111/phpr.12136

Burge, T. (2022). *Perception : First Form of Mind*. Oxford University Press.

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, *45*(4), 431-448. https://doi.org/10.1016/j.visres.2004.08.021

Butler, A., Oruc, I., Fox, C. J., & Barton, J. J. S. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Research*, *1191*, 116-126. https://doi.org/10.1016/j.brainres.2007.10.101

Butterfill, S. (2009). Seeing Causings and Hearing Gestures. *The Philosophical Quarterly (1950-)*, *59*(236), 405-428.

Calzavarini, F., & Voltolini, A. (2022). Perception of Faces and Other Progressively Higher-Order Properties. *Topoi*, *41*(4), 671-684. https://doi.org/10.1007/s11245-022-09802-4

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision : The first 1000 ms. *Journal of Vision*, *13*(10), 1-1. https://doi.org/10.1167/13.10.1

Carranante, G. (2020). *The Boundaries of Perception : A Conceptual Engineering Pespective*. Université Paris Sciences et Lettres.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7(3), 308-313. https://doi.org/10.1038/nn1194

Casale, M. B., & Ashby, F. G. (2008). A Role for the Perceptual Representation Memory System in Category Learning. *Perception & psychophysics*, *70*(6), 983-999.

Cavanagh, P. (2011). Visual cognition. *Vision Research*, *51*(13), 1538-1551. https://doi.org/10.1016/j.visres.2011.01.015

Cave, C. B. (1997). Very Long-Lasting Priming in Picture Naming. *Psychological Science*, *8*(4), 322-325. https://doi.org/10.1111/j.1467-9280.1997.tb00446.x

Cavedon-Taylor, D. (2021). High-Level Perception and Multimodal Perception. In H. Logue & L. Richardson (Éds.), *Purpose and Procedure in Philosophy of Perception* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198853534.003.0008

Chalmers, D. J. (2004). The Representational Character of Experience. In B. Leiter (Éd.), *The Future for Philosophy* (p. 153-181). Oxford University Press.

Chalmers, D. J. (2006). Perception and the Fall From Eden. In T. S. Gendler & J. Hawthorne (Éds.), *Perceptual Experience* (p. 49-125). Oxford University Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55-81. https://doi.org/10.1016/0010-0285(73)90004-2

Cheung, O. S., & Bar, M. (2014). The resilience of object predictions : Early recognition across viewpoints and exemplars. *Psychonomic Bulletin & Review*, *21*(3), 682-688. https://doi.org/10.3758/s13423-013-0546-5

Chudnoff, E. (2011). The Nature of Intuitive Justification. *Philosophical Studies*, *153*(2), 313-333. https://doi.org/10.1007/s11098-010-9495-2

Chudnoff, E. (2020). *Forming Impressions : Expertise in Perception and Intuition*. Oxford University Press.

Churchland, P. M. (1979). *Scientific Realism and the Plasticity of Mind*. Cambridge University Press. https://doi.org/10.1017/CBO9780511625435

Churchland, P. M. (1988). Perceptual Plasticity and Theoretical Neutrality : A Reply to Jerry Fodor. *Philosophy of Science*, *55*(2), 167-187. JSTOR.

Churchland, P. M. (2012). *Plato's camera : How the physical brain captures a landscape of abstract universals*. MIT Press.

Clark, A. (1992). *Sensory Qualities*. Clarendon Press.

Coggan, D. D., Liu, W., Baker, D. H., & Andrews, T. J. (2016). Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *NeuroImage*, *135*, 107-114. https://doi.org/10.1016/j.neuroimage.2016.04.060

Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., Abdi, H., & Haxby, J. V. (2012). The Representation of Biological Classes in the Human Brain. *Journal of Neuroscience*, *32*(8), 2608-2618. https://doi.org/10.1523/JNEUROSCI.5547-11.2012

Connolly, K. (2014). Perceptual Learning and the Contents of Perception. *Erkenntnis*, *79*(6), 1407-1418. https://doi.org/10.1007/s10670-014-9608-y

Connolly, K. (2019). *Perceptual learning : The flexibility of the senses*. Oxford University Press.

Contini, E. W., Wardle, S. G., & Carlson, T. A. (2017). Decoding the time-course of object recognition in the human brain : From visual features to categorical decisions. *Neuropsychologia*, *105*, 165-176. https://doi.org/10.1016/j.neuropsychologia.2017.02.013

Cooke, S. F., & Bear, M. F. (2015). Visual recognition memory : A view from V1. *Current opinion in neurobiology*, *35*, 57-65. https://doi.org/10.1016/j.conb.2015.06.008

Cooper, L. A., Schacter, D. L., Ballesteros, S., & Moore, C. (1992). Priming and recognition of transformed three-dimensional objects : Effects of size and reflection. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*(1), 43-57. https://doi.org/10.1037//0278-7393.18.1.43

Cowan, R. (2015). Cognitive penetrability and ethical perception. *Review of Philosophy and Psychology*, *6*(4), 665-682. https://doi.org/10.1007/s13164-014-0185-4

Cox, D. D. (2014). Do we understand high-level vision? *Current Opinion in Neurobiology*, *25*, 187-193. https://doi.org/10.1016/j.conb.2014.01.016

Crane, T. (2009). Is Perception a Propositional Attitude? *The Philosophical Quarterly (1950-)*, *59*(236), 452-469.

Crouzet, S., & Thorpe, S. (2011). Low-Level Cues and Ultra-Fast Face Detection. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00342

Cummins, R. (2000). Reply to Millikan. *Philosophical and Phenomenological Research*, *60*(1), 113-127. https://doi.org/10.2307/2653431

Davies, W. (2018). Colour Vision and Seeing Colours. *British Journal for the Philosophy of Science*, *69*(3), 657-690. https://doi.org/10.1093/bjps/axw026

De Brigard, F. (2014). The Nature of Memory Traces. *Philosophy Compass*, *9*. https://doi.org/10.1111/phc3.12133

Delk, J. L., & Fillenbaum, S. (1965). Differences in Perceived Color as a Function of Characteristic Color. *The American Journal of Psychology*, *78*(2), 290-293. https://doi.org/10.2307/1420503

Deroy, O. (2013). Object-sensitivity versus cognitive penetrability of perception. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *162*(1), 87-107.

Deroy, O. (2019). Categorising without Concepts. *Review of Philosophy and Psychology*, *10*(3), 465-478. https://doi.org/10.1007/s13164-019-00431-2

Devereux, B. J., Clarke, A., & Tyler, L. K. (2018). Integrated deep visual and semantic attractor neural networks predict fMRI pattern-information along the ventral object processing pathway. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-28865-1

de Vignemont, F. (2023). Fifty Shades of Affective Colouring of Perception. *Australasian Journal of Philosophy*, *101*(1), 1-15. https://doi.org/10.1080/00048402.2021.1965176

Diamond, R., & Carey, S. (1986). Why faces are and are not special : An effect of expertise. *Journal of Experimental Psychology. General*, *115*(2), 107-117. https://doi.org/10.1037//0096-3445.115.2.107

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333-341. https://doi.org/10.1016/j.tics.2007.06.010

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415-434. https://doi.org/10.1016/j.neuron.2012.01.010

Dokic, J. (2010). Perceptual Recognition and the Feeling of Presence. In B. Nanay (Éd.), *Perceiving the World* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195386196.003.0003

Dokic, J., & Martin, J.-R. (2015). 'Looks the Same but Feels Different' : A Metacognitive Approach to Cognitive Penetrability. In *The Cognitive Penetrability of Perception : New Philosophical Perspectives*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198738916.001.0001/acprof-9780198738916-chapter-11

Dretske, F. (1969). *Seeing and knowing*. University of Chicago Press.

Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press.

Dretske, F. (1986). Misrepresentation. In R. Bogdan (Éd.), *Belief : Form, Content, and Function* (p. 17-36). Oxford University Press.

Dretske, F. (1995). *Naturalizing the mind*. MIT Press.

Dretske, F. (2015). Perception versus Conception : The Goldilocks Test. In *The Cognitive Penetrability of Perception : New Philosophical Perspectives*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198738916.001.0001/acprof-9780198738916-chapter-7

Dubova, M., & Goldstone, R. L. (2021). The Influences of Category Learning on Perceptual Reconstructions. *Cognitive Science*, *45*(5), e12981. https://doi.org/10.1111/cogs.12981

Edelman, S. (1999). *Representation and recognition in vision*. MIT Press.

Firestone, C., & Scholl, B. (2016a). Seeing and thinking : Foundational issues and empirical horizons. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X16000029

Firestone, C., & Scholl, B. J. (2015). When do ratings implicate perception versus judgment? The "overgeneralization test" for top-down effects. *Visual Cognition*, *23*(9-10), 1217-1226. https://doi.org/10.1080/13506285.2016.1160171

Firestone, C., & Scholl, B. J. (2016b). Cognition does not affect perception : Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X15000965

Fish, W. (2013). High-level properties and visual experience. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *162*(1), 43-55.

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1338-1349. https://doi.org/10.1098/rstb.2011.0417

Fodor, J. A. (1983). *The Modularity of Mind : An Essay on Faculty Psychology*. A Bradford Book / MIT Press.

Fodor, J. A. (1984). Observation Reconsidered. *Philosophy of Science*, *51*(1), 23-43. JSTOR.

Fodor, J. A. (1987). *Psychosemantics : The problem of meaning in the philosophy of mind*. MIT Press.

Fodor, J. A. (1998). *Concepts : Where Cognitive Science Went Wrong* (1st edition). Clarendon Press.

Fodor, J. A. (2007). The Revenge of the Given. In B. P. McLaughlin & J. Cohen (Éds.), *Contemporary Debates in Philosophy of Mind*. Blackwell. https://philpapers.org/rec/FODTRO

Gärdenfors, P. (2000). *Conceptual spaces : The geometry of thought*. Cambridge, MA. : MIT Press. http://archive.org/details/conceptualspaces0000gard

Gauker, C. (2011). *Words and images : An essay on the origin of ideas*. Oxford University Press.

Gauker, C. (2017). Three Kinds of Nonconceptual Seeing-as. *Review of Philosophy and Psychology*, *8*(4), 763-779. https://doi.org/10.1007/s13164-017-0339-2

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), Article 4. https://doi.org/10.1038/nn1029

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" Expert : Exploring Mechanisms for Face Recognition. *Vision Research*, *37*(12), 1673-1682. https://doi.org/10.1016/S0042-6989(96)00286-6

Gauthier, I., & Tarr, M. J. (2016). Visual Object Recognition : Do We (Finally) Know More Now Than We Did? *Annual Review of Vision Science*, *2*, 377-396. https://doi.org/10.1146/annurev-vision-111815-114621

Gaziv, G., Lee, M. J., & DiCarlo, J. J. (2023). *Robustified ANNs Reveal Wormholes Between Human Category Percepts* (arXiv:2308.06887). arXiv. https://doi.org/10.48550/arXiv.2308.06887

Gelman, S. A. (2003). *The Essential Child : Origins of Essentialism in Everyday Thought*. OUP USA.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *8*(9), 404-409. https://doi.org/10.1016/j.tics.2004.07.001

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

Goldman, A. I. (1976). Discrimination and Perceptual Knowledge. *The Journal of Philosophy*, *73*(20), 771-791. https://doi.org/10.2307/2025679

Goldstone, R. L. (1994a). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology. General*, *123*(2), 178-200. https://doi.org/10.1037//0096-3445.123.2.178

Goldstone, R. L. (1994b). The role of similarity in categorization : Providing a groundwork. *Cognition*, *52*(2), 125-157. https://doi.org/10.1016/0010-0277(94)90065-5

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, *65*(2), 231-262. https://doi.org/10.1016/S0010-0277(97)00047-4

Goldstone, R. L., & Byrge, L. A. (2015). Perceptual learning. In *The Oxford handbook of philosophy of perception* (p. 814-832). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199600472.001.0001

Goldstone, R. L., Landy, D., & Brunel, L. (2011). Improving Perception to Make Distant Connections Closer. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00385

Goodman, N. (1972). Seven Strictures on Similarity. In *Problems and Projects*. Bobs-Merril.

Green, E. J. (2017). Psychosemantics and the Rich/Thin Debate. *Philosophical Perspectives*, *31*(1), 153-186. https://doi.org/10.1111/phpe.12097

Green, E. J. (2020a). The Perception-Cognition Border : A Case for Architectural Division. *The Philosophical Review*, *129*(3), 323-393. https://doi.org/10.1215/00318108-8311221

Green, E. J. (2020b). *The Perception-Cognition Border : A Case for Architectural Division*. https://philpapers.org/rec/GRETPB-5

Greene, M. R., & Oliva, A. (2009). The briefest of glances : The time course of natural scene understanding. *Psychological Science*, *20*(4), 464-472. https://doi.org/10.1111/j.1467-9280.2009.02316.x

Greene, M. R., & Oliva, A. (2010). High-level aftereffects to global scene properties. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1430-1442. https://doi.org/10.1037/a0019058

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition : As soon as you know it is there, you know what it is. *Psychological Science*, *16*(2), 152-160. https://doi.org/10.1111/j.0956-7976.2005.00796.x

Gross, S. (2017). Cognitive Penetration and Attention. *Frontiers in Psychology*, *8*, 221. https://doi.org/10.3389/fpsyg.2017.00221

Gross, S., Chaisilprungraung, T., Kaplan, E., Menendez, J. A., & Flombaum, J. (2014). Problems for the Purported Cognitive Penetration of Perceptual Color Experience and Macpherson's Proposed Mechanism. In E. Machery & J. Prinz (Éds.), *Baltic International Yearbook of Cognition, Logic and Communication* (p. 01-30).

Hadjikhani, N., Kveraga, K., Naik, P., & Ahlfors, S. P. (2009). Early (N170) activation of face-specific cortex by face-like objects. *Neuroreport*, *20*(4), 403-407. https://doi.org/10.1097/WNR.0b013e328325a8e1

Hafri, A., Landau, B., Bonner, M. F., & Firestone, C. (2019). When a phone in a basket looks like a knife in a cup : Perception and abstraction of visual-spatial relations between objects. *Journal of Vision*, *19*(10), 160a. https://doi.org/10.1167/19.10.160a

Hahn, U., Prat-Sala, M., Pothos, E. M., & Brumby, D. P. (2010). Exemplar similarity and rule application. *Cognition*, *114*(1), 1-18. https://doi.org/10.1016/j.cognition.2009.08.011

Hahn, U., & Ramscar, M. (Éds.). (2001). *Similarity and categorization*. Oxford University Press.

Hansen, M. K. (2018). Perception of High-Level Content and the Argument from Associative Agnosia. *Review of Philosophy and Psychology*, *9*(2), 301-312. https://doi.org/10.1007/s13164-017-0364-1

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, *9*(11), 1367-1368. https://doi.org/10.1038/nn1794

Helton, G. (2016). Recent issues in high-level perception. *Philosophy Compass*, *11*(12), 851-862. https://doi.org/10.1111/phc3.12383

Hepper, P. (1996). Fetal memory : Does it exist? What does it do? *Acta Paediatrica*, *85*(s416), 16-20. https://doi.org/10.1111/j.1651-2227.1996.tb14272.x

Hillinger, M. L. (1980). Priming effects with phonemically similar words: *Memory & Cognition*, *8*(2), 115-123. https://doi.org/10.3758/BF03213414

Hohwy, J. (2013). *The Predictive Mind*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199682737.001.0001/acprof-9780199682737

Holzinger, Y., Ullman, S., Harari, D., Behrmann, M., & Avidan, G. (2019). Minimal Recognizable Configurations Elicit Category-selective Responses in Higher Order Visual Cortex. *Journal of Cognitive Neuroscience*, *31*(9), 1354-1367. https://doi.org/10.1162/jocn_a_01420

Horgan, T., & Tienson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In D. J. Chalmers (Éd.), *Philosophy of Mind : Classical and Contemporary Readings* (p. 520-533). Oup Usa.

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210-1224. https://doi.org/10.1016/j.neuron.2012.10.014

Javadi, A. H., & Wee, N. (2012). Cross-Category Adaptation : Objects Produce Gender Adaptation in the Perception of Faces. *PLOS ONE*, *7*(9), e46079. https://doi.org/10.1371/journal.pone.0046079

Johnson, D. J., & Wilson, J. P. (2019). Racial Bias in Perceptions of Size and Strength : The Impact of Stereotypes and Group Differences. *Psychological Science*, *30*(4), 553-562. https://doi.org/10.1177/0956797619827529

Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams : Recalling the past and imagining the future. *Science*, *367*(6473), eaaw4325. https://doi.org/10.1126/science.aaw4325

Kanizsa, G. (1985). Seeing and thinking. *Acta Psychologica*, *59*(1), 23-33. https://doi.org/10.1016/0001-6918(85)90040-X

Keil, F. C. (1992). *Concepts, Kinds, and Cognitive Development*. https://doi.org/10.7551/mitpress/2065.001.0001

Kelly, S. D. (2001). Demonstrative Concepts and Experience. *The Philosophical Review*, *110*(3), 397-420. https://doi.org/10.2307/2693650

Kim, J. S., Aheimer, B., Montané Manrara, V., & Bedny, M. (2021). Shared understanding of color among sighted and blind adults. *Proceedings of the National Academy of Sciences*, *118*(33), e2020192118. https://doi.org/10.1073/pnas.2020192118

Kim, J. S., Elli, G. V., & Bedny, M. (2019). Knowledge of animal appearance among sighted and blind adults. *Proceedings of the National Academy of Sciences*, *116*(23), 11213-11222. https://doi.org/10.1073/pnas.1900952116

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements : Visual processing speed revisited. *Vision Research*, *46*(11), 1762-1776. https://doi.org/10.1016/j.visres.2005.10.002

Kosakowski, H. L., Cohen, M. A., Takahashi, A., Keil, B., Kanwisher, N., & Saxe, R. (2022). Selective responses to faces, scenes, and bodies in the ventral visual pathway of infants. *Current Biology*, *32*(2), 265-274.e5. https://doi.org/10.1016/j.cub.2021.10.064

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The Case for Mental Imagery*. https://doi.org/10.1093/acprof:oso/9780195179088.001.0001

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway : An expanded neural framework for the processing of object quality. *Trends in cognitive sciences*, *17*(1), 26-49. https://doi.org/10.1016/j.tics.2012.10.011

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry : Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401-412. https://doi.org/10.1016/j.tics.2013.06.007

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*. https://www.frontiersin.org/articles/10.3389/neuro.06.004.2008

Kulvicki, J. V. (2004). Isomorphism in Information Carrying Systems. *Pacific Philosophical Quarterly*, *85*(4), 380-395. https://doi.org/10.1111/j.1468-0114.2004.00206.x

Lammers, N. A., de Haan, E. H., & Pinto, Y. (2017). No Evidence of Narrowly Defined Cognitive Penetrability in Unambiguous Vision. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00852

Lande, K. J. (2018). The Perspectival Character of Perception. *The Journal of Philosophy*, *115*(4), 187-214. https://doi.org/10.5840/jphil2018115413

Lande, K. J. (In Press). Pictorial syntax. *Mind & Language*. https://doi.org/10.1111/mila.12497

Landers, C. (2021). Specialized Visual Experiences. *Philosophical Quarterly*, *71*(1), 74-98. https://doi.org/10.1093/pq/pqaa018

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*(1), Article 1. https://doi.org/10.1038/82947

Levine, D. N. (1978). Prosopagnosia and visual object agnosia : A behavioral study. *Brain and Language*, *5*(3), 341-365. https://doi.org/10.1016/0093-934X(78)90031-7

Levine, J. (2010). Demonstrative Thought. *Mind and Language*, *25*(2), 169-195. https://doi.org/10.1111/j.1468-0017.2009.01385.x

Lewis, D. (1980). Veridical hallucination and prosthetic vision. *Australasian Journal of Philosophy*, *58*(3), 239-249. https://doi.org/10.1080/00048408012341251

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596-9601. https://doi.org/10.1073/pnas.092277599

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing : Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron*, *62*(2), 281-290. https://doi.org/10.1016/j.neuron.2009.02.025

Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, *17*(6), 20. https://doi.org/10.1167/17.6.20

Luck, S. J., & Hollingworth, A. (Éds.). (2008). *Visual Memory*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305487.001.0001

Lupyan, G. (2015). Cognitive Penetrability of Perception in the Age of Prediction : Predictive Systems are Penetrable Systems. *Review of Philosophy and Psychology*, *6*(4), 547-569. https://doi.org/10.1007/s13164-015-0253-4

Lycan, W. G. (2014). What Does Vision Represent? In *Does Perception Have Content?* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199756018.003.0013

Lyons, J. C. (2005). Clades, Capgras, and Perceptual Kinds. *Philosophical Topics*, *33*(1), 185-206. https://doi.org/10.5840/philtopics20053317

Lyons, J. C. (2022). Three grades of iconicity in perception. *Asian Journal of Philosophy*. https://philpapers.org/rec/LYOTGO

Machery, E. (2009). *Doing Without Concepts*. New York, US: Oxford University Press.

Machery, E. (2016). The amodal brain and the offloading hypothesis. *Psychonomic Bulletin & Review*, *23*(4), 1090-1095. https://doi.org/10.3758/s13423-015-0878-4

MacPherson, F. (2006). Ambiguous Figures and the Content of Experience. *Noûs*, *40*(1), 82-117.

Macpherson, F. (2012). Cognitive Penetration of Colour Experience : Rethinking the Issue in Light of an Indirect Mechanism. *Philosophy and Phenomenological Research*, *84*(1), 24-62. https://doi.org/10.1111/j.1933-1592.2010.00481.x

Madole, K. L., & Oakes, L. M. (1999). Making Sense of Infant Categorization : Stable Processes and Changing Representations. *Developmental Review*, *19*(2), 263-296. https://doi.org/10.1006/drev.1998.0481

Mandelbaum, E. (2015). The automatic and the ballistic : Modularity beyond perceptual processes. *Philosophical Psychology*, *28*(8), 1147-1156. https://doi.org/10.1080/09515089.2014.950217

Mandler, J. M. (2000). Perceptual and Conceptual Processes in Infancy. *Journal of Cognition and Development*, *1*(1), 3-36. https://doi.org/10.1207/S15327647JCD0101N_2

Maniglia, M., & Seitz, A. R. (2018). Towards a whole brain model of Perceptual Learning. *Current Opinion in Behavioral Sciences*, *20*, 47-55. https://doi.org/10.1016/j.cobeha.2017.10.004

Marotta, J. J., McKeeff, T. J., & Behrmann, M. (2002). The effects of rotation and inversion on face processing in prosopagnosia. *Cognitive Neuropsychology*, *19*(1), 31-47. https://doi.org/10.1080/02643290143000079

Marr, D. (1982). *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H.Freeman & Co Ltd.

Martin, C. B., & Barense, M. D. (2023). Perception and Memory in the Ventral Visual Stream and Medial Temporal Lobe. *Annual Review of Vision Science*, *9*, 409-434. https://doi.org/10.1146/annurev-vision-120222-014200

Martin, C. B., Douglas, D., Newsome, R. N., Man, L. L., & Barense, M. D. (2018). Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *eLife*, 7, e31873. https://doi.org/10.7554/eLife.31873

Matsumiya, K. (2013). Seeing a haptically explored face : Visual facial-expression aftereffect from haptic adaptation to a face. *Psychological Science*, *24*(10), 2088-2098. https://doi.org/10.1177/0956797613486981

Matthen, M. (2007). *Seeing, Doing, and Knowing : A Philosophical Theory of Sense Perception*. Oxford University Press.

McDowell, J. (1994). *Mind and World*. Harvard University Press.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254-278. https://doi.org/10.1037/0033-295X.100.2.254

Mednick, S. A., & Freedman, J. L. (1960). Stimulus generalization. *Psychological Bulletin*, *57*(3), 169-200. https://doi.org/10.1037/h0041650

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7), 2. https://doi.org/10.1167/4.7.2

Michaelian, K. (2011). Generative Memory. *Philosophical Psychology*, *24*(3), 323-342. https://doi.org/10.1080/09515089.2011.559623

Michotte, A. (1963). *The perception of causality* (p. xxii, 424). Basic Books.

Millar, A. (2000). The Scope of Perceptual Knowledge. *Philosophy*, *75*(291), 73-88.

Millikan, R. (1989). Biosemantics. In B. P. McLaughlin & A. Beckerman (Éds.), *Journal of Philosophy* (p. 281-297). Oxford University Press.

Morrison, J. (2016). Perceptual Confidence. *Analytic Philosophy*, *57*(1), 15-48. https://doi.org/10.1111/phib.12077

Moutsopoulou, K., Pfeuffer, C., Kiesel, A., Yang, Q., & Waszak, F. (2019). How long is long-term priming? Classification and action priming in the scale of days. *Quarterly Journal of Experimental Psychology (2006)*, *72*(5), 1183-1199. https://doi.org/10.1177/1747021818784261

Munton, J. (2019). Perceptual Skill And Social Structure. *Philosophy and Phenomenological Research*, *99*(1), 131-161. https://doi.org/10.1111/phpr.12478

Murphy, G. L. (2004). *The big book of concepts* (1. MIT Press paperback ed). MIT Press.

Murphy, J., Gray, K. L. H., & Cook, R. (2017). The composite face illusion. *Psychonomic Bulletin & Review*, *24*(2), 245-261. https://doi.org/10.3758/s13423-016-1131-5

Nanay, B. (2011). Do We See Apples as Edible? *Pacific Philosophical Quarterly*, *92*(3), 305-322. https://doi.org/10.1111/j.1468-0114.2011.01398.x

Nanay, B. (2013). *Between Perception and Action*. Oxford University Press.

Navarro, V. M., Jani, R., & Wasserman, E. A. (2019). Pigeon Category Learning : Revisiting the Tasks. *Journal of experimental psychology. Animal learning and cognition*, *45*(2), 174-184. https://doi.org/10.1037/xan0000198

Neander, K. (2017). *A Mark of the Mental : In Defense of Informational Teleosemantics*. MIT Press.

Nes, A., Sundberg, K., & Watzl, S. (2023). The perception/cognition distinction. *Inquiry*, *66*(2), 165-195. https://doi.org/10.1080/0020174X.2021.1926317

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(42), 16598-16603. https://doi.org/10.1073/pnas.0703913104

Newen, A. (2017). Defending the liberal-content view of perceptual experience : Direct social perception of emotions and person impressions. *Synthese*, *194*(3), 761-785. https://doi.org/10.1007/s11229-016-1030-3

Nida-Rümelin, M. (1996). Pseudonormal Vision : An Actual Case of Qualia Inversion? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *82*(2), 145-157.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57. https://doi.org/10.1037/0096-3445.115.1.39

Oliva, A., & Torralba, A. (2001). Modeling the Shape of the Scene : A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*, *42*(3), 145-175. https://doi.org/10.1023/A:1011139631724

Openshaw, J. (2022). Remembering Objects. *Philosophers' Imprint*, *22*(n/a), 1-20. https://doi.org/10.3998/phimp.699

Orlandi, N. (2011). The Innocent Eye : Seeing-as Without Concepts. *American Philosophical Quarterly*, *48*(1), 17-31.

Overgaard, S. (2022). Seeing-as, seeing-o, and seeing-that. *Philosophical Studies*, *179*, 1-20. https://doi.org/10.1007/s11098-022-01810-9

Pacherie, E. (2005). *Perceiving intentions.*

Palmer, S. E. (1999). *Vision science : Photons to phenomenology*. MIT Press.

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In *Attention and Performance IX* (p. 135-151). Erlbaum.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *23*(2), 324-354. https://doi.org/10.1037//0278-7393.23.2.324

Palmeri, T. J., & Tarr, M. J. (2008). Visual Object Perception and Long-term Memory. In S. J. Luck & A. Hollingworth (Éds.), *Visual Memory* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305487.003.0006

Papineau, D. (2021). *The Metaphysics of Sensory Experience*. Oxford University Press. https://doi.org/10.1093/oso/9780198862390.001.0001

Pautz, A. (2009). A Simple View of Consciousness. In R. C. Koons & G. Bealer (Éds.), *The Waning of Materialism* (p. 25-66). Oxford University Press. https://philarchive.org/rec/PAUCAS

Peacocke, C. (1983). *Sense and Content : Experience, Thought and their Relations*. Oxford University Press.

Pelser, A. (2010). Belief in Reid's Theory of Perception. *History of Philosophy Quarterly*, *27*(4), 359-378.

Perky, C. W. (1910). An Experimental Study of Imagination. *The American Journal of Psychology*, *21*(3), 422-452. https://doi.org/10.2307/1413350

Peters, B., & Kriegeskorte, N. (2021). Capturing the objects of vision with neural networks. *Nature Human Behaviour*, *5*(9), Article 9. https://doi.org/10.1038/s41562-021-01194-6

Phillips, B. (2019). The Shifting Border Between Perception and Cognition. *Noûs*, *53*(2), 316-346. https://doi.org/10.1111/nous.12218

Pitcher, D., Goldhaber, T., Duchaine, B., Walsh, V., & Kanwisher, N. (2012). Two Critical and Functionally Distinct Stages of Face and Body Perception. *The Journal of Neuroscience*, *32*(45), 15877-15885. https://doi.org/10.1523/JNEUROSCI.2624-12.2012

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353-363. https://doi.org/10.1037/h0025953

Price, R. (2009a). Aspect-Switching and Visual Phenomenal Character. *The Philosophical Quarterly (1950-)*, *59*(236), 508-518.

Price, R. (2009b). *The Ways Things Look*. University of Oxford.

Prinz, J. J. (2002). *Furnishing the Mind : Concepts and Their Perceptual Basis*. MIT Press.

Prinz, J. J. (2006). Beyond Appearances : The Content of Sensation and Perception. *Perceptual Experience*. https://doi.org/10.1093/acprof:oso/9780199289769.003.0013

Prinz, J. J. (2012). *The conscious brain : How attention engenders experience*. Oxford University Press.

Prinz, J. J. (2013). Siegel's get rich quick scheme. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *163*(3), 827-835.

Proklova, D., Kaiser, D., & Peelen, M. V. (2016). Disentangling representations of object shape and object category in human visual cortex : The animate–inanimate distinction. *Journal of Cognitive Neuroscience*, *28*(5), 680-692. https://doi.org/10.1162/jocn_a_00924

Pryor, J. (2000). The Skeptic and the Dogmatist. *Noûs*, *34*(4), 517-549.

Putnam, H. (1975). The Meaning of « Meaning ». *Minnesota Studies in the Philosophy of Science*, 7, 131-193.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *The Behavioral and Brain Sciences*, *22*(3), 341-365; discussion 366-423. https://doi.org/10.1017/s0140525x99002022

Quilty-Dunn, J. (2017). Syntax and Semantics of Perceptual Representation. *Dissertations, Theses, and Capstone Projects*. https://academicworks.cuny.edu/gc_etds/2261

Quilty-Dunn, J. (2019). Reid's Doxastic Theory of Perception. *Philosophers' Imprint*, *19*.

Quilty-Dunn, J. (2021). Polysemy and Thought : Toward a Generative Theory of Concepts. *Mind and Language*, *36*(1), 158-185. https://doi.org/10.1111/mila.12328

Quine, W. V. (1969). Natural Kinds. In J. Kim & E. Sosa (Éds.), *Ontological Relativity and Other Essays* (p. 114-138). Columbia University Press.

Quinn, P. C., & Eimas, P. D. (1997). A Reexamination of the Perceptual-to-Conceptual Shift in Mental Representations. *Review of General Psychology*, *1*(3), 271-287. https://doi.org/10.1037/1089-2680.1.3.271

Quinn, P. C., & Eimas, P. D. (2000). The emergence of category representations during infancy : Are separate and conceptual processes required? *Journal of Cognition and Development*, *1*(1), 55-61. https://doi.org/10.1207/S15327647JCD0101N_6

Quinn, P. C., Eimas, P. D., & Tarr, M. J. (2001). Perceptual categorization of cat and dog silhouettes by 3- to 4-month-old infants. *Journal of Experimental Child Psychology*, *79*(1), 78-94. https://doi.org/10.1006/jecp.2000.2609

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42-55. https://doi.org/10.1038/nrn.2016.150

Ramsey, W. (2015). Untangling two questions about mental representation. *New Ideas in Psychology*, *40*. https://doi.org/10.1016/j.newideapsych.2015.01.004

Ransom, M. (2019). *Perceptual learning for expertise* [University of British Columbia]. https://doi.org/10.14288/1.0389682

Ransom, M. (2020a). Attentional Weighting in Perceptual Learning. *Journal of Consciousness Studies*, *27*(7-8), 236-248.

Ransom, M. (2020b). Expert Knowledge by Perception. *Philosophy*, *95*(3), 309-335. https://doi.org/10.1017/S0031819120000157

Ransom, M. (2020c). Waltonian Perceptualism. *The Journal of Aesthetics and Art Criticism*, *78*(1), 66-70. https://doi.org/10.1111/jaac.12685

Ratcliff, G., & Newcombe, F. (1982). Object recognition : Some deductions from the clinical evidence. In A. W. Ellis (Éd.), *Normality and pathology in cognitive functions*. Academic Press.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998). Contrasting Cortical Activity Associated with Category Memory and Recognition Memory. *Learning & Memory*, *5*(6), 420-428.

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Current Biology*, *27*(12), 1825-1828.e3. https://doi.org/10.1016/j.cub.2017.05.044

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To See or not to See : The Need for Attention to Perceive Changes in Scenes. *Psychological Science*, *8*(5), 368-373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Riddoch, M. J., & Humphreys, G. W. (1987a). A case of integrative visual agnosia. *Brain: A Journal of Neurology*, *110 (Pt 6)*, 1431‑1462. https://doi.org/10.1093/brain/110.6.1431

Riddoch, M. J., & Humphreys, G. W. (1987b). Visual object processing in optic aphasia : A case of semantic access agnosia. *Cognitive Neuropsychology*, *4*(2), 131‑185. https://doi.org/10.1080/02643298708252038

Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, *12*(2), 162‑168. https://doi.org/10.1016/S0959‑4388(02)00304‑5

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences*, *14*(2), 64‑71. https://doi.org/10.1016/j.tics.2009.11.005

Romanski, L. M., & Chafee, M. V. (2021). A View from the Top : Prefrontal Control of Object Recognition. *Neuron*, *109*(1), 6‑8. https://doi.org/10.1016/j.neuron.2020.12.014

Rosch, E., & Mervis, C. B. (1975). Family resemblances : Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573‑605. https://doi.org/10.1016/0010‑0285(75)90024‑9

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision*, *8*(12), 3. https://doi.org/10.1167/8.12.3

Rubens, A. B., & Benson, D. F. (1971). Associative visual agnosia. *Archives of Neurology*, *24*(4), 305‑316. https://doi.org/10.1001/archneur.1971.00480340037003

Sacchi, E. (2011). A phenomenologically oriented account of the phenomenon of aspectuality in propositional attitudes. *Phenomenology and Mind*, *1*, Article 1. https://doi.org/10.13128/Phe_Mi‑19644

Sacchi, E. (2018). Fregean Presentationalism. In A. Coliva, P. Leonardi, & S. Moruzzi (Éds.), *Eva Picardi on Language, Analysis and History* (p. 241‑261). Springer International Publishing. https://doi.org/10.1007/978‑3‑319‑95777‑7_11

Sacchi, E., & Voltolini, A. (2017). Against Phenomenal Externalism. *Crítica (México D. F. En Línea)*, *49*(145), 25‑48. https://doi.org/10.22201/iifs.18704905e.2017.195

Sacks, O. (1998). *The Man Who Mistook His Wife For A Hat : And Other Clinical Tales*. Touchstone.

Saysani, A., Corballis, M. C., & Corballis, P. M. (2021). Seeing colour through language : Colour knowledge in the blind and sighted. *Visual Cognition*, *29*(1), 63‑71. https://doi.org/10.1080/13506285.2020.1866726

Schacter, D. L. (1990). Perceptual representation systems and implicit memory : Toward a resolution of the multiple memory systems debate. *Annals of the New York Academy of Sciences*, *608*, 543‑571. https://doi.org/10.1111/j.1749‑6632.1990.tb48909.x

Schindler, A., & Bartels, A. (2016). Visual high‑level regions respond to high‑level stimulus content in the absence of low‑level confounds. *NeuroImage*, *132*, 520‑525. https://doi.org/10.1016/j.neuroimage.2016.03.011

Scholl, B., & Gao, T. (2013). Perceiving animacy and intentionality : Visual processing or higher‑level judgment? *Social perception: Detection and interpretation of animacy, agency, and intention*. https://doi.org/10.7551/mitpress/9780262019279.003.0009

Schulte, P. (2012). *How Frogs See the World : Putting Millikan's Teleosemantics to the Test*. https://philpapers.org/rec/SCHHFS

Schulte, P. (2021). The Nature of Perceptual Constancies. *Philosophy and Phenomenological Research*, *103*(1), 3‑20. https://doi.org/10.1111/phpr.12693

Schulte, P., & Neander, K. (2022). Teleological Theories of Mental Content. In E. N. Zalta (Éd.), *The Stanford Encyclopedia of Philosophy* (Summer 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/content-teleological/

Schurgin, M. W. (2018). Visual memory, the long and the short of it : A review of visual working memory and long-term memory. *Attention, Perception, & Psychophysics*, *80*(5), 1035-1056. https://doi.org/10.3758/s13414-018-1522-y

Schweinberger, S. R., Zäske, R., Walther, C., Golle, J., Kovács, G., & Wiese, H. (2010). Young without plastic surgery : Perceptual adaptation to the age of female and male faces. *Vision Research*, *50*(23), 2570-2576. https://doi.org/10.1016/j.visres.2010.08.017

Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *The Behavioral and Brain Sciences*, *21*(1), 1-17; discussion 17-54. https://doi.org/10.1017/s0140525x98000107

Sellars, W. (1978). The Role of the Imagination in Kant's Theory of Experience. In H. W. Johnstone (Éd.), *Categories : A Colloquium*. Pennsylvania State University Press.

Serre, T. (2016). Models of visual categorization. *WIREs Cognitive Science*, 7(3), 197-213. https://doi.org/10.1002/wcs.1385

Shallice, T., & Cooper, R. P. (2013). Is there a semantic system for abstract words? *Frontiers in Human Neuroscience*, 7. https://doi.org/10.3389/fnhum.2013.00175

Shea, N. (2014). Exploitable Isomorphism and Structural Representation. *Proceedings of the Aristotelian Society*, *114*, 123-144.

Shea, N. (2018). *Representation in Cognitive Science*. Oxford University Press. https://doi.org/10.1093/oso/9780198812883.001.0001

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations : Shapes of states. *Cognitive Psychology*, *1*(1), 1-17. https://doi.org/10.1016/0010-0285(70)90002-2

Shimamura, A. P. (1986). Priming effects of amnesia : Evidence for a dissociable memory function. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *38*(4), 619-644. https://doi.org/10.1080/14640748608401617

Shoemaker, S. (2003). Content, Character and Color. *Philosophical Issues*, *13*(1), 253-278. https://doi.org/10.1111/1533-6077.00014

Siegel, S. (2006). Which Properties are Represented in Perception? In *Perceptual Experience*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199289769.003.0015

Siegel, S. (2011). *The Contents of Visual Experience*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195305296.001.0001/acprof-9780195305296

Siegel, S. (2013a). Replies to Campbell, Prinz, and Travis. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *163*(3), 847-865.

Siegel, S. (2013b). The epistemic impact of the etiology of experience. *Philosophical Studies*, *162*(3), 697-722. https://doi.org/10.1007/s11098-012-0059-5

Siegel, S. (2017). *The Rationality of Perception*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198797081.001.0001/acprof-9780198797081

Siegel, S., & Byrne, A. (2017). Rich or Thin? In B. Nanay (Éd.), *Current Controversies in Philosophy of Perception*. New York, USA: Routledge.

Siewert, C. P. (1998). *The significance of consciousness*. Princeton University Press.

Skrzypulec, B. zej. (2018). Perceptual Kinds as Supervening Sortals. *Pacific Philosophical Quarterly*, *100*(1), 174-201. https://doi.org/10.1111/papq.12253

Sloman, S. A., & Rips, L. J. (1998). Similarity as an explanatory construct. *Cognition*, *65*(2), 87-101. https://doi.org/10.1016/S0010-0277(97)00048-6

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When Looks Are Everything : Appearance Similarity versus Kind Information in Early Induction. *Psychological Science*, *18*(2), 179-185.

Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, *65*(2-3), 167-196. https://doi.org/10.1016/s0010-0277(97)00043-7

Smith, J. D., Zakrzewski, A. C., Johnson, J. M., Valleau, J. C., & Church, B. A. (2016). Categorization : The View from Animal Cognition. *Behavioral Sciences*, *6*(2), Article 2. https://doi.org/10.3390/bs6020012

Smith, L. B., & Heise, D. (1992). Perceptual Similarity and Conceptual Structure. In B. Burns (Éd.), *Advances in Psychology* (Vol. 93, p. 233-272). North-Holland. https://doi.org/10.1016/S0166-4115(08)61009-2

Smortchkova, J. (2021). After-effects and the reach of perceptual content. *Synthese*, *198*(8), 7871-7890. https://doi.org/10.1007/s11229-020-02554-x

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. *Proceedings of the National Academy of Sciences*, *119*(43), e2200800119. https://doi.org/10.1073/pnas.2200800119

Soto, F. A., & Wasserman, E. A. (2014). Mechanisms of object recognition : What we have learned from pigeons. *Frontiers in Neural Circuits*, *8*. https://www.frontiersin.org/articles/10.3389/fncir.2014.00122

Spoerer, C. J., McClure, P., & Kriegeskorte, N. (2017). Recurrent Convolutional Neural Networks : A Better Model of Biological Object Recognition. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01551

Stokes, D. (2015). Towards a Consequentialist Understanding of Cognitive Penetration. In *The Cognitive Penetrability of Perception*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198738916.003.0003

Stokes, D. (2018a). Attention and the Cognitive Penetrability of Perception. *Australasian Journal of Philosophy*, *96*(2), 303-318. https://doi.org/10.1080/00048402.2017.1332080

Stokes, D. (2018b). Rich Perceptual Content and Aesthetic Properties. In A. Bergqvist & R. Cowan (Éds.), *Evaluative Perception*. Oxford University Press.

Stokes, D. (2021a). On Perceptual Expertise. *Mind and Language*, *36*(2), 241-263. https://doi.org/10.1111/mila.12270

Stokes, D. (2021b). *Thinking and perceiving*. Routledge.

Storrs, K. R. (2015). Are high-level aftereffects perceptual? *Frontiers in Psychology*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00157

Strawson, P. F. (2008). *Freedom and resentment and other essays*. Routledge.

Taylor, H. (2020). Fuzziness in the Mind : Can Perception be Unconscious? *Philosophy and Phenomenological Research*, *101*(2), 383‑398. https://doi.org/10.1111/phpr.12592

Thompson, B. (2007a). Shoemaker on Phenomenal Content. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *135*(3), 307‑334.

Thompson, B. (2007b). Shoemaker on Phenomenal Content. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *135*(3), 307‑334.

Thompson, B. (2009). Senses for Senses. *Australasian Journal of Philosophy*, *87*(1), 99‑117. https://doi.org/10.1080/00048400802215471

Thompson, B. (2010). The Spatial Content of Experience. *Philosophy and Phenomenological Research*, *81*(1), 146‑184.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520‑522. https://doi.org/10.1038/381520a0

Tiedemann, H., Morgenstern, Y., Schmidt, F., & Fleming, R. W. (2022). One‑shot generalization in humans revealed through a drawing task. *eLife*, *11*, e75485. https://doi.org/10.7554/eLife.75485

Tranel, D., & Damasio, A. R. (1988). Non‑conscious face recognition in patients with face agnosia. *Behavioural Brain Research*, *30*(3), 235‑249. https://doi.org/10.1016/0166‑4328(88)90166‑0

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301‑306. https://doi.org/10.1126/science.2296719

Tye, M. (1995). *Ten Problems of Consciousness : A Representational Theory of the Phenomenal Mind*. MIT Press.

Ullman, S., Assif, L., Fetaya, E., & Harari, D. (2016). Atoms of recognition in human and computer vision. *Proceedings of the National Academy of Sciences*, *113*(10), 2744‑2749. https://doi.org/10.1073/pnas.1513198113

Valenti, J. J., & Firestone, C. (2019). Finding the « odd one out » : Memory color effects and the logic of appearance. *Cognition*, *191*, 103934. https://doi.org/10.1016/j.cognition.2019.04.003

VanRullen, R. (2006). On second glance : Still no high‑level pop‑out effect for faces. *Vision Research*, *46*(18), 3017‑3027. https://doi.org/10.1016/j.visres.2005.07.009

VanRullen, R., & Thorpe, S. J. (2001). Is it a Bird? Is it a Plane? Ultra‑Rapid Visual Categorisation of Natural and Artifactual Objects. *Perception*, *30*(6), 655‑668. https://doi.org/10.1068/p3029

Voltolini, A. (2023). Perceiving Aesthetic Properties. *The British Journal of Aesthetics*, *63*(3), 417‑434. https://doi.org/10.1093/aesthj/ayac039

Voss, J. L., Baym, C. L., & Paller, K. A. (2008). Accurate forced‑choice recognition without awareness of memory retrieval. *Learning & Memory*, *15*(6), 454‑459. https://doi.org/10.1101/lm.971208

Wagemans, J. (Éd.). (2015). *The Oxford handbook of perceptual organization* (First edition). Oxford University Press.

Walton, K. L. (1970). Categories of Art. *The Philosophical Review*, *79*(3), 334‑367. https://doi.org/10.2307/2183933

Wapner, W., Judd, T., & Gardner, H. (1978). Visual agnosia in an artist. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *14*(3), 343‑364. https://doi.org/10.1016/s0010‑9452(78)80062‑8

Watson, D. M., Young, A. W., & Andrews, T. J. (2016). Spatial properties of objects predict patterns of neural response in the ventral visual pathway. *NeuroImage*, *126*, 173‑183. https://doi.org/10.1016/j.neuroimage.2015.11.043

Watzl, S. (2017). *Structuring Mind : The Nature of Attention and how it Shapes Consciousness*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199658428.001.0001

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*(6982), 557‑561. https://doi.org/10.1038/nature02420

Weiskopf, D. A. (2015). Observational Concepts. In E. Margolis & S. Laurence (Éds.), *The Conceptual Mind : New Directions in the Study of Concepts*. MIT Press.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*(2), 227‑233. https://doi.org/10.1016/s0959-4388(98)80144-x

Wimmer, M. C., Doherty, M. J., & Collins, W. A. (2011). The Development of Ambiguous Figure Perception. *Monographs of the Society for Research in Child Development*, *76*(1), i‑130.

Wittgenstein, L., Hacker, P. M. S., & Schulte, J. (2009). *Philosophical Investigations.* (4e édition). Wiley-Blackwell.

Wu, W. (2017). Shaking Up the Mind's Ground Floor : The Cognitive Penetration of Visual Attention. *Journal of Philosophy*, *114*(1), 5‑32. https://doi.org/10.5840/jphil201711411

Xu, X., Yue, X., Lescroart, M. D., Biederman, I., & Kim, J. G. (2009). Adaptation in the fusiform face area (FFA) : Image or person? *Vision Research*, *49*(23), 2800‑2807. https://doi.org/10.1016/j.visres.2009.08.021

Zachariou, V., Del Giacco, A. C., Ungerleider, L. G., & Yue, X. (2018). Bottom-up processing of curvilinear visual features is sufficient for animate/inanimate object categorization. *Journal of Vision*, *18*(12), 3. https://doi.org/10.1167/18.12.3

Zeimbekis, J. (2013). Color and cognitive penetrability. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *165*(1), 167‑175. JSTOR.

Zeman, A. A., Ritchie, J. B., Bracci, S., & Op de Beeck, H. (2020). Orthogonal Representations of Object Shape and Category in Deep Convolutional Neural Networks and Human Visual Cortex. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-59175-0

Zhao, M., & Hayward, W. G. (2010). Holistic processing underlies gender judgments of faces. *Attention, Perception, & Psychophysics*, *72*(3), 591‑596. https://doi.org/10.3758/APP.72.3.591