Framing Effects in Object Perception

Spencer Ivy and Aleksandra Mroczko-Wąsowicz*

Department of Philosophy, University of Warsaw

Forthcoming in Review of Philosophy and Psychology

Abstract

In this paper we argue that object perception may be affected by what we call "perceptual frames." Perceptual frames are adaptations of the perceptual system that guide how perceptual objects are singled out from a sensory environment. These adaptations are caused by perceptual learning and realized through bottom-up functional processes such that sensory information is organized in a subject-dependent way leading to idiosyncratic perceptual object representations. Through domain-specific training, perceptual learning, and the acquisition of object-knowledge, it is possible to modulate the adaptive perceptual system such that its ability to represent becomes bespoke. Different perceivers with different perceptual frames may, therefore, receive the same sensory information and perceive different perceptual objects due to the effects of framing. Consequently, we demonstrate the plausibility of this account by surveying empirical data concerning the functions of (1) multisensory integration, (2) amodal completion, and (3) predictive anticipation. Regarding (1), we argue that the perceptual system's optimization processes employ perceptual frames to facilitate multisensory feature binding. Regarding (2), we argue that amodal completion can occur with or without the help of mental imagery, yet either instance of amodal completion requires perceptual frames. Regarding (3), we demonstrate that perceptually anticipating an object's motion involves the implementation of perceptual frames. We conclude that framing effects are a matter of perceptual diversity and highlight the need to accommodate unique perspectives in the philosophy and science of perception.

Keywords: object perception; cognition; predictive processing; Bayesianism; multisensory integration; amodal completion; mental imagery; anticipation; top-down; bottom-up

* Correspondence: aleksandra.mroczko-wasowicz@uw.edu.pl, University of Warsaw, Faculty of Philosophy, Krakowskie Przedmieście 3, 00-097 Warszawa, Poland

I. Introduction

Imagine, for a moment, the following scenario. It's Friday night and you and your friends are going out to the local symphony hall for an evening of chamber music. Tonight is something of a special occasion because not only are you going to hear your favorite piece of music performed – Bach's Brandenburg Concerto No. 2 – but you are also bringing along a special guest. A friend has invited a professor of music from the nearby university who specializes in Bach to join you on your evening out. The performance is wonderful and after it's all over, you discuss your experience with your friends. As a fan of Bach, you can hold your own in levying aesthetic judgments as you are a reliably accurate listener who was priorly aware of the piece. Nevertheless, you are amazed at the level of description in detail and complexity offered by the music professor. You are especially amazed at discovering that the professor could hear Bach's name spelled by the composition of the piece.[1] You sat next to this person in the symphony hall, you listened to the same music being played, but you both experienced it very differently. Did you hear the same piece of music?

Of course, the simplest answer to the foregoing question is yes: you didn't receive any different perceptual information than the Bach scholar who sat next to you. But still, the depth of your own experience seems to have paled in comparison to that of the expert. One way to immediately account for these differences would be to conjecture that with a richer conceptual apparatus and prior knowledge of what to expect (and, in general, the effects of scholarly expertise), the professor was much better cognitively equipped to analyze and describe the received auditory information than you and your friends. While this explanation is very likely true, in the paper to follow we argue for an additional explanation. Despite having sat through the very same symphony, what is actually perceived will vary due to the effects of cognition and attention on how perceptual information is received, processed, and ultimately organized in experience. Thus, the expert really did hear something different than you and your friends due to the unique way in which their expertise caused them to attend the symphony.

[1] This is a figure of composition known as the BACH motif where notes are composed from corresponding letters of the alphabet that, when arranged, can spell Bach's name. It was also famously used by Shostakovich in modes of both love and political protest.

As we shall demonstrate, the effects of cognition on perception are not merely a matter of post-perceptually conceptualizing and organizing perceptual information. We argue that cognition can genuinely change what perceptual information is possible to be perceived and how it may be represented. We argue that cognitive processes can affect the formation of perceptual experience by providing or limiting accessibility to possible perceptual objects. We call the cause of these effects "Perceptual Frames" and show how they constitute evidence of genuine interactions across the joint of cognition and perception. To preview, a perceptual frame is an adaptation of the perceptual system realized through bottom-up functional processes, typically developed via perceptual learning, that changes representations given in object perception. So, for example, the Bach scholar will have experienced a unique representation of the motif within the symphony precisely because their adapted perceptual system has learned to individuate and represent the motif's features as a unique perceptual object. What causes the individuation of the motif as a perceptual object is attributed to the perceptual system and its adaptations rather than to the direct, top-down influence of the Bach scholar's knowledge; at least, so we shall argue in the remainder of this paper.

Furthermore, we argue that perceptual frames differ from competing accounts of mental imagery and attention. As we demonstrate, the perceptual system is adaptive, and its adaptations – perceptual frames – are not merely matters of attention, perspective, or imagery, but rather are normalized and conditioned background functions which govern how sensory information is processed. Accordingly, we neither assume that cognition penetrates perception, nor do we take a stance on whether or not perceptual content is rich, high level, or conceptual. Instead, for the sake of argument, we assume a conservative approach concerning the joint between cognition and perception. In doing so, we show that even in the case that perception and cognition are encapsulated from one another, there are nevertheless non-trivial effects from cognition on perception.[2]

[2] While we assume this conservative approach for the sake of our argument, our view may still yet accommodate the possibility of cognitive penetration, admissible contents of perception, and other claims to cognition's effect on perception. What we argue says nothing about the effects that cognition may or may not have on the representation of perceptual features. Rather, what we argue for here is that cognition plays an important role at the level of establishing what counts as a perceptual object.

To support these claims, in the following two sections of this paper we define in further detail what perceptual frames are and how they are comparable to a broad category of cognitive and attentional processes directed at the perceptual processing of incoming sensory signals. Following this, in section three we survey a selection of empirical research demonstrating that these perceiver-specific frames can drive or limit multimodal and amodal forms of perceptual completion and object individuation. We conclude in the final section of the paper that whatever metaphysical conditions and constraints are proffered to explain perceptual object individuation, we ought not to ignore what the perceiver themselves bring to bear in the construction of perceptual objects.

II. Different perspectives, different perceptions - commonalities and differences between perceptual frames and top-down attentional processes

Perceptual objects (also sometimes called 'sensory individuals,' 'objects for perception,' or 'object files') are the basic building blocks upon which our perceptual capacities operate to represent the world (Mroczko-Wąsowicz and Grush 2023). Although there are important technical and philosophical distinctions between all three of these terms, for the sake of this paper, we will consider them all to share the (hopefully) uncontroversial commonality of being a singular target for perception. This minimal definition is what we mean to express by "perceptual object." As Casey O'Callaghan (2016: pg. 1270) so neatly says: "in perception, objects are key." The perceptual system interacts with the world by dividing it up into objects and events and attributing features to them. These feature-attributed-individuals are then represented in the perceiver's (typically) conscious experience. Accordingly, perceptual objects are the individuals to whom represented features are attributed. Importantly, these processes of individuation and feature-attribution are governed by rules concerning what counts as a possible perceptual object and how features are bound (or not) to their objects. We shall call these rules the 'structural principles' of perceptual objects.

According to most any account of perceptual objects and their individuation, the perceptual system identifies individuals and represents their features under the guidance of basic structural principles (Mroczko-Wąsowicz, Stoch and Zguda, 2023). "These principles specify which arrangements of parts count as objects for the relevant perceptual system, but also ensure that many arrangements do not count" (Green, 2018: pg. 665). What basic structural principles count as fundamental for perception varies by just about any given view. Examples of these principles can include shared temporal and spatial location, cohesion, connectedness, boundedness, Gestalt laws, and the list goes on (*See* Wagemans et al., 2012; Casati, 2015 *for an overview of some of the more commonly accepted principles*). Likewise, O'Callaghan (2008, 2016) and Cohen (2023) have argued that perceptual objects are 'mereological complexes' and Green (2018) has provided a convincing account of causal regularity as a structural principle that is essential to perceptual object individuation. In any case, whatever structural principles that your favorite account of perceptual objecthood entails, the fundamental story is the same: perception works by systematically individuating objects in the environment.

In addition, object individuation may be modulated by perspectival aspects of sense perception. This concerns the ways that a perceptual object with the same sensory properties may be experienced as appearing differently depending on the context or on the perceiving subject (Green and Schellenberg, 2018; Schwenkler, 2014). The latter, interestingly, emphasizes the importance of the causal influence of perceptual relations (or the constitutive role of a subject) on the status of a perceptual object. Thus, as Skrzypulec (2022: pg. 2840-2841) puts it, "The existence of subject-dependent perceptual objects is a result of the proper functioning of perceptual mechanisms that organize the visual scene according to the cognitive interests of a subject." Skrzypulec' point, here, is important. He argues that attention-driven relations between the perceiver and their environment are to be counted as what we have named above, a 'structural principle' of perceptual object individuation. In other words, perceptual objects aren't always perceiver-independent individuals nebulously waiting within their environments to be represented. Rather, *how* a perceiver identifies, reidentifies, and generally attributes perceptual features to a perceptual scene can perform a necessary perspectival function in constituting perceptual objects within that scene.

In this manner of speaking, perceptual objects can be subject-dependent, which, for Skrzypulec means occupying a particular attentional perspective (Skrzypulec 2023; see also Alsmith 2017). Without such a perspective, as Skrzypulec argues, there is no singularly principled way in which certain environmental perceptual features may be bound together. Accordingly, the account of perceptual frames that we forward in the next section of the paper makes a similar claim about the subject-dependency of perceptual objects, yet does not rely upon top-down shifts of attention to determine subject-dependent object individuation. However, to motivate this claim, it is helpful to first see how attentional accounts of subject-dependency work before setting them aside to investigate a further claim about the effects of perceptual frames on bottom-up sensory processing.

Skrzypulec focuses on perceivers' 'attentional interests' as constitutive of a perspective that enables the individuation of perceptual objects. He asks his readers to consider a lattice of six identical black dots. To shift one's attention to different sets of dots, he argues, will constitute the individuation of distinct perceptual objects because each set of dots will accord with what we have called the structural principles of object perception. Any two-dot segment can stand apart as a figure from the ground of the lattice, the two dots compose a distinct mereological structure, and the features of any set of two dots will remain unchanged through time (*ibid*). These structural principles of perceptual object individuation are satisfied by the two-dot combinations *and* which two-dot combinations become perceptual objects depends on the perspectival direction of attention. Significantly, all of this perceptual object individuation can be done through 'covert' shifts in attention (Chalmers, 2004; Block, 2010). That is to say, the manipulations can be performed without moving one's eyes and collecting any new perceptual information. Therefore, without a change to the inputs of perception, the same perceptual scene can afford different sets of perceptual objects given the perspectival character of attention.

Similar arguments to Skrzypulec' remain a growing source of interest in psychological research on the relationship between attention and the representation of perceptual objects. Ongchoco & Scholl (2022, 2023) have introduced the phenomenon of "scaffolded attention;" namely, that people may perceive structured objects and events that do not arise directly from sensory cues. For instance, when staring at a regular grid or listening to a regular beat, subjects may perceive emergent spatial or temporal structures (see Figure 1c). Meng & Tong (2006) similarly demonstrated how the top-down direction of covert attention is strongly correlated with changes in how Necker cubes are perceptually experienced (see Figure 1a). Additionally, Carrasco (2004, 2006, 2018) has demonstrated and argued that manipulations of attention can cause changes in how contrast, brightness, and the directional orientation of perceptual objects are consciously represented. Thus, attention has a part to play in how perceptual objects are individuated from their environment and whether any set of features is collected together to be attributed to a perceptual object.

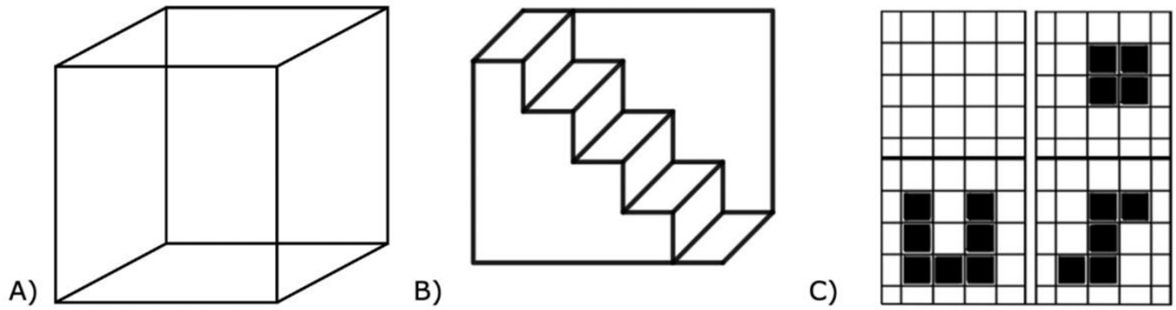

Fig 1. Each of these images are examples of bistable or attention-driven changes to representation. The Necker cube (a)[3] can be shifted from top to bottom-up views. The Schroeder stairs (b)[4] can be represented as normal or inverted. And, the final image (c) presents three different structured patterns resulting from the scaffolded attention for different perceivers viewing a blank grid.

That being said, it seems possible that there are some objects for which attention (or perceptual frames) do not play a necessarily constitutive role. Some objects may be simple enough to have their necessary and sufficient conditions met by the basic structural principles of perception. For instance, any perceiver with normally functioning perceptual capacities should be able to individuate any of the six dots in Skrzypulec' lattice on the basis of their color, arrangement, and orientation. This is to say, then, that not all perceptual objects are necessarily subject-dependent – especially those that are sufficiently simple material objects for which any perceiver with regularly functioning perceptual capacities will be likely to perceive as such (Matthen, 2023). Yet, it is equally possible that another class of perceptual objects remains subject-dependent precisely because of the perspectival character of perception. How and where attention is directed is one such way in which different perceivers might individuate different perceptual objects given the same perceptual environment and sensory inputs.

However, what unifies these attentional accounts of perspectival object perception is that they (1) describe subject-dependent representation on the basis of top-down direction; and (2) yield

[3] https://en.wikipedia.org/wiki/Necker_cube
[4] https://en.wikipedia.org/wiki/Schroeder_stairs

perspectives that are universally accessible to any perceiver with normally functioning capacities. Yet, as we demonstrate in what follows, idiosyncratic and non-universal perceptual object representations are possible without top-down directed shifts in attention. The preference for specific perceptual object individuation can be accommodated by the perceptual system before attention is applicable. Consequently, the ways in which perceptual objects are individuated on a subject-dependent basis *before* the application of attention shall be our focus for the remainder of this paper. That is, our interest regards how perceptual processing can be influenced in a subject-dependent manner through strictly bottom-up processing of sensory information.

III. The Perceptual frame and its effects on object individuation

What the foregoing accounts show is that for any set of simple perceptual features, there are different perceptual objects available to access from distinct attentional perspectives. For simple tasks like differentiating dots in a lattice, shifting the contrast of overlapping circles, or switching the view of bistable images, any perceiver with normally functioning perceptual capacities should be able to perform the requisite manipulation to find the desired perspective. Anyone can be directed to 'see the Necker cube from a top/bottom view' then shift their perspective to the alternative view; and, anyone can select two from a set of six dots as the focus of their attention to find in them a unique perceptual object. There are, however, some perspectives (like those of the perceptual expert) that are unavailable to every perceiver with normally functioning capacities – perspectives within which unique perceptual objects can be perceived. Not everyone can identify the sufficient set of notes constitutive of the Bach motif apart from the rest of the concerto as its own, individual object of perception. The perspective from which it would be possible to individuate the motif's notes as a distinct perceptual object is one that requires training and a specific coordination of one's perceptual capacities. Accordingly, we shall call this perceiver-specific coordination of perceptual capacities the "perceptual frame." We define perceptual frames as follows:

> **Perceptual Frame:** A perceptual frame is an adaptation of the perceptual system caused by perceptual learning and realized through bottom-up functional processes such that sensory information is organized in a subject-dependent way leading to idiosyncratic perceptual object representations.

Given the foregoing definition, a perceptual frame is a kind of perspective like those discussed above with regard to top-down shifts of attention. Yet, not all perspectives require the focus of attention, nor a perceiver's interests, nor occurrent top-down influences in order to function. There are also *default* perspectives which live in the background and are simply the adaptively conditioned ways that we perceive in a given environment. These default perspectives function as a heuristic for perceptual processing and can be trained, developed, or adapted to prefer the selection and representation of objects[5] in a sensory environment. Moreover, there is no such thing as a ubiquitous perspective shared by all perceivers – everyone is a little (or a lot) different.

For example, perceptual experts experience unique affordances and search strategies for their domain of expertise (Kundel et al., 2007; Gegenfurtner et al., 2011; Sheridan and Reingold, 2017; Stokes, 2021; Ivy et al., 2021; Ivy 2023). The ability to see more completely and more efficiently than novices is a hallmark of perceptual training and expertise. One could account for this kind of effect by invoking the concept of "mental imagery". Mental imagery is a form of perceptual completion that does not depend on sensory stimulation, but rather familiarity with what perceptual objects should be (Nanay 2010; 2023). For instance, knowing which stars make up a constellation, one can apply a mental image of that constellation to see those stars as an object apart from the rest of the night sky (Briscoe, 2011). The application of mental imagery to a sensory scene can be automatic and also subject-dependent. Two perceivers may know different constellations and so see different combinations in the night sky. Thus, expert perceivers may simply apply a more rigorous or complete mental image to better perceive the scenes that they are familiar with. To the contrary, there is empirical evidence that the subject-dependent effects on expert object perception come prior to the application of mental images to familiar scenes.

Eye-tracking studies have shown that expert-specific search strategies are applied (though not effectively) to search tasks outside of the experts' domain; non-experts do not exhibit the same distinctive search patterns, nor do they search as accurately or efficiently as experts in domain-oriented images (Ivy et al., 2023). Therefore, we may interpret these findings as demonstrating

[5] Although our argument here focuses on framing effects for objects, it seems plausible as well that such effects can be attributed to features. For instance, the salience of a color's contrast may be increased by a perceptual frame causing that color to pop out relative to an object's other features. This effect would not change anything structural about the object besides how its colors are represented with respect to one another, yet would still indicate the influence of a perceptual frame. Thanks to an anonymous reviewer for pointing this out.

that perceptual experts have a default mode of perceiving applied to both their domain of expertise and outside of it. Further, the default mode of perceiving is specific to each type of expert inasmuch as different domains of expertise require different search strategies for success. The way that an expert moves their eyes has been deliberately trained and developed through time to respond to the perceptual objects within their domain. Although these defaulted strategies bleed over into other tasks, it is only for an expert's own domain that their perceptual frame helps to organize the sensory scene accurately and efficiently.

If it were the case that perceptual experts applied mental images to efficiently search and represent sensory scenes, we would expect them to only apply those mental images where they are applicable – within their domains of expertise. You don't see constellations where there are no stars. However, given that experts employed the same search strategies both in and outside of their domains of expertise, this seems not to be the case. The way that experts are trained to look, they indiscriminately apply while searching any scene (but are only better off for it when searching within their domain). This does not mean that, for example, radiologists look for tumors in the cars lining the highway on their daily commute. Rather, it means that the unique saccadic pattern that they have developed to expertly find tumors in radiographs are vestigial behaviors present in the way that they gaze at the highway while looking for opportunities to pass (among other search tasks). Put into the terms of the present paper's thesis, the perceptual frame is the search pattern that operates in a bottom-up manner indiscriminate of the target task or goal; however, when the target task or scene match the perceptually framed search pattern, tumors pop out for the radiologist as they do not for those with untrained eyes.

Perhaps there is still something to be said for mental imagery regarding the functions of expert perception, but the foregoing evidence suggests that there are other pre-perceptual, (and thus, pre-imagistic) subject-dependent forces at play; i.e., perceptual frames. Expert perceptual capacities are trained and developed, and thus causally connected to rich cognitive apparatus aimed at adapting to the targets and goals of search. Yet, like a motorbike whose engine and chassis have been built for the racetrack, the bike can still operate inefficiently on the street even if it was not designed for that purpose. Similarly, perceptually framed adaptations within the perceptual system will prefer the sensory environments that they are attuned to regardless of attention or mental imagery. Or in other words, the specific way that the motorbike is designed for the racetrack (the perceptual frame) does not bear upon how the bike is being ridden (the application of mental imagery) when considering what kinds of maneuvers the bike is able to perform.

This is what makes a perceptual frame a unique kind of perspective. How perceptual information is received and arranged can be trained and developed. However, top-down influences are not required at the time of perceiving in order for sensory information to be organized by the defaulted perceptual frame. Acute, expert forms of perception can function automatically; and insofar as they are default heuristics of the perceptual system, they apply to scenes ubiquitously. In this manner of speaking, the perceptual frame operates as a kind of attunement to particular kinds of sensory information as well as to a specific directive for sensory information to be organized by the perceptual system. This is what enables radiologists to accurately diagnose images flashed before them at incredibly fast speeds up to 200 ms/image (Drew et al., 2013). The radiologist's perceptual frame is attuned to automatically favor particular sensory inputs associated with what they know aberrancies to typically look like, and where they are most likely to be found (Haider & Frensch, 1999; Brams et al., 2019). However, there is an important question yet to be answered: can the organization and sensitivity to sensory information afforded by a perceptual frame be sufficient for the individuation of perceptual objects which would otherwise be inaccessible without that frame?

To answer yes to this question amounts to saying that it is possible for any two perceivers with different perceptual frames to have distinctly different perceptual experiences given the same sensory input. Moreover, if it is the case that from the same set of sensory information two perceivers may have the ability to individuate unique perceptual objects inaccessible to the other perceiver, then the effect of perceptual frames on perceptual processing is both deep and pervasive. Where the structure of perceptual processing and experience is contingent upon a perceptual frame, that frame operates akin to the structural principles of perceptual objecthood discussed in section one. Without the appropriate frame, possible perceptual objects could not be individuated or identified. Accordingly, in the three subsections to follow, we survey a series of empirical data from perceptual processing in multisensory, amodal, and predictive settings to support the foregoing suppositions. We conclude on the basis of this evidence that perceptual frames often play a necessary role for perceptual objecthood.

i. Bayesian processing and multisensory integration

There is a conceptual problem known by many names in the philosophy and cognitive science of perception: *the binding problem* (Treisman, 1998), *the many properties problem* (Jackson 1977; Clark 2000), *the many-many problem*[6] (Wu, 2014), *the ambiguity of sensory combination* (Ernst & Bülthoff, 2004), etc. The problem is that in any sensory environment there are infinitely many possible combinations of sensory features. The question is then raised - how are our perceptual capacities able to discriminate, integrate, and circumscribe multitudes of sensory data into a coherent and veridical representation of the environment? Given what we have reported above, a partial answer to this question can be offered in the form of the structural principles of perceptual objecthood. The structural principles of perceptual objecthood sufficiently explain the circumscription and integration of basic sensory features into simple perceptual objects. However, as the sensory environment becomes either more complex or more ambiguous, the structural principles cease to sufficiently explain feature integration. This is why, as we argued above, perceptual frames make a significant impact – especially for trained and expert perceivers.

The problem becomes much more of a challenge when we take into account the fact that the sensory environment rarely, if ever, is entirely unisensory. Just about every perceptual scene that we encounter affords sensory information about the same objects processed by different sense modalities. We can hear a guitar play just as well as we can feel its strings when plucked, or see our fingers set against its frets. Given that sensory environments contain multiple sources of sensory information about the same perceptual objects, many of those information streams are redundant upon one another. Yet, despite complex sensory noise, perceptual ambiguity, and multisensory redundancy, the perceptual system is quite adept at representing clear, integrated, and coherently structured environments. As we contend in what follows, this is possible because perceptual frames often help to make these adjudications. This fact is exemplified by an influential account of Bayesian perceptual processing called the Maximum Likelihood Estimation (MLE).

---

[6] The "many-many" problem is slightly different from the others that are listed insofar as it relates to agentive control in cases where agents face too many perceptual inputs and too many possible behavioral outputs. However, we believe that the intuition behind Wu's many-many problem and its argumentative form run in parallel to the others. For example, if perceptual objects are actually event files that include action-consequences, then the many-many problem is a kind of binding problem with respect to which action-consequence type of information is bound in the event file (Mroczko-Wąsowicz et al., 2024).

MLE holds that redundant streams of perceptual information are integrated into a single multisensory representation if and only if doing so optimizes veridical representations of perceptual objects. If it is the case that integrating a guitar's auditory properties with its visual and tactile properties will afford an optimal representation (most veridical) of that guitar, then the perceptual system will do so. However, where redundant sensory data may be noisy or imperfect, MLE predicts that the perceptual system will rely upon other more valid sensory data to individuate perceptual objects. Significantly, the success of such calculations requires a sensitivity to what sensory information counts as noise, what counts as valid, and what sorts of combinations will yield optimal representations. Given that this sensitivity affects bottom-up sensory processing, is trainable and makes a significant difference in the representation of perceptual objects, it is a strong candidate for a perceptual frame.

For example, in one experiment (Ernst & Banks, 2002), subjects were asked to determine the height of a bar that could be both felt and seen. When the visual information was muddied with extra visual noise, the subjects relied on touch to make their judgments; the opposite effect was found when tactile inputs were made noisier than the visual inputs. Further, by varying the degree of noise across either visual or tactile inputs, the researchers developed a Bayesian model that is able to predict *to what degree* any stream of sensory information will be utilized by the perceptual system when individuating objects. According to the model, if two streams of sensory data are redundant and equally reliable (e.g., you can both clearly see and feel an object), both sets of data will be used equally to make estimations about the object. However, when noise is added to a redundant sensory input, the perceptual system weights that noise against the input's reliability for constructing an optimal perceptual representation. Thus, the perceptual system will rely more upon the clear sensory data and less upon the noisy sensory data inversely proportional to the amount of noise present (see Figure 2 presenting the optimizing functions of the perceptual system as demonstrated in a pair of studies by Kirsch and Kunde, 2023).

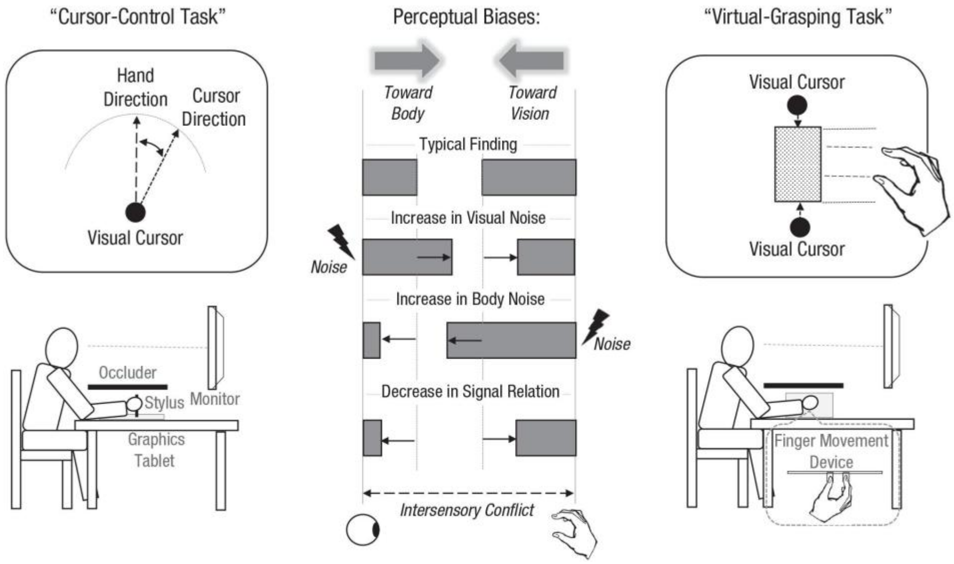

Fig 2. In the first task on the left, participants judged the distance that a target moves on a screen by tracing a stylus below an occluder. In the second task on the right, participants place marks on either side of an object to enclose it. In both tasks visual and tactile noise were introduced to demonstrate how the participants' perceptual systems responded to bias by preferring the alternative, optimal stream of sensory information. The middle graphic demonstrates this optimizing strategy in different instantiations of intersensory conflict. From Kirsch and Kunde (2023), used under Creative Commons CC BY license.

Thus, MLE's Bayesian model predicts how the perceptual system utilizes sensory information as a function of that information's probability to produce optimally accurate perceptual object representations. This explains, in part, how the perceptual system discriminates between different streams of sensory information to integrate, bind, and individuate perceptual objects with accurate reliability. Thus, at this point, we have the data we need to begin to make a case for perceptual frames on the basis of multisensory feature integration. Inasmuch as the perceptual system reliably assigns values of probability to multiple streams of sensory information in order to construct an optimal estimation of the perceptual scene, *this* is the beginning of a perceptual frame. The sensory information that is counted as clear is integrated, and the perceptual information that is counted as noisy is left out. Further, the principle by which the perceptual system is ready to assign these probabilities is latently active, ready to be applied in the act of perceiving.

Significantly, it is important to note that the original and basic formulations of MLE's Bayesian function made consistent predictions across study participants on the assumption that normalized assignments of probability are baked into normally functioning perceptual capacities.

That is to say, the original calculations of MLE do not account for differences in the assignment of probability, or, in other words – a difference of perceptual frame. If the basic probability schema counts as a perceptual frame, it is a frame shared by the vast majority of perceivers. Nevertheless, MLE's model can account for unique perceptual frames bespoke to perceivers whose perceptual systems integrate multisensory information differently. Recent developments in MLE that have focused on "bias" and "prior knowledge" do just this (Helbig and Ernst, 2007; Ernst, 2012; Mandrigin, 2018). These developments indicate that where there is "prior knowledge," there is a perceptual frame at work.

It is not uncommon for a perceptual system to be biased by producing non-veridical representations of the world. For instance, without glasses, one's eyes may represent the world as blurrier than with glasses, and similarly – tinnitus can cause one to hear ringing where none exists. Extreme cases of perceptual bias have long been an area of interest for researchers who have developed studies that manipulate vision by distorting the world 180° vertically (Stratton, 1896, 1897; Helmholtz, 1924), or some small degree to the left or right (Held, 1961; Held and Bossom, 1961). What the data of these studies show is that over time, the perceptual system will re-tool itself to adjust to the biased shift of sensory presentation. Whereas one might overreach for an object due to their vision being off by 35°, after training, coordination eventually returns. The perceptual system is capable of learning the difference between what is veridical and what is the bias through which it represents the world. This difference is prior knowledge, and likewise an example of a perceptual frame.

In cases where perceptual information is potentially biased, "this bias may be unstable, for example because of fast adaptation processes that constantly react to small discrepancies… The brain could learn this bias uncertainty and use this knowledge to emphasize the more stable estimates" (Ernst and Bülthoff, 2004: pg. 168). In other words, the brain and perceptual system can learn to discriminate perceptual information that has been biased by noise from optimally veridical representations. On the basis of this learned perceptual information, bias can be taken into account when calculating the maximum likelihood of an object's veridicality (Helbig & Ernst, 2007; Ernst, 2012). Significantly, because the processes and goals of perceptual learning can differ between observers, so too can the consequent calculations that are made to integrate reliable multisensory cues into perceptual objects. Accordingly, the perceptual systems of different perceivers who bias perceptual information differently may, in fact, produce unique perceptual objects given the same sensory inputs. Insofar as a perceiver's prior knowledge is uniquely attuned to their own history of perceptual learning and overcoming of perceptual bias, the basic functions of their multisensory integration will differ from those with a different history. Thus, the effects of perceptual frames are demonstrated by the perceptual system's learned response to bias.

For instance, in studies that introduce bias by shifting visual perception askew, adaptation effects are the measured results of subjects who remove the vision-shifting glasses and perceive the ordinary world as if it were still offset (e.g., the world is no longer skewed 15 degrees to the right, but subjects still mis-reach for objects). Significantly, only the study participants who actively learned to adapt to the skewed vision have adverse adaptation effects after taking off the world-shifting glasses (*See* Bermejo et al., 2020 *for a review*). Subjects that did not interact with the surroundings during the time of wearing the glasses do not have adaptation effects. That is, participants who controlled their movement actively adapted to the skewed perceptual world by implementing a bias-responsive perceptual frame which persisted even after taking the glasses off. In contrast, participants who did not interact with the environment did not adapt their perceptual frame to the new visual information. Accordingly, when they removed the glasses, they slipped right back into their normally-tracking perceptual frame. Therefore, these studies are evidence that perceptual frames are learned adaptations of the perceptual system.[7] Although the frames are constructed with the aid of top-down influence (i.e., subjects needed to be in control of their movements in order to adapt), their operations occur bottom-up (i.e., adaptation effects occurred only for the group that constructed the bias-responsive perceptual frame after taking off the glasses).

To bring it all together, MLE is a Bayesian model that predicts whether features from multiple sense modalities are integrated into a single multisensory object, and if so – how. The principal calculation predicted by MLE is a function of the perceptual system to seek and to

[7] Expert perceivers also apply their bias-responsive perceptual strategies to search tasks both in and out of their domains of expertise, even though they are only better perceivers for their own domain (Ivy et al., 2023). This is because the gaze pattern that experts utilize for successful search within their domain is a part of their perceptual frame. These frames are successful for experts because they account for bias and prior knowledge specific to perceptual targets of their domain of expertise. Outside of their domain of expertise, the bias-overcoming strategies and prior knowledge no longer track relevant objects in the sensory scene and so fail to increase accuracy or efficiency of search.

organize sensory environments such that they are optimal: maximally veridical and minimally variant, noisy, or biased. Prior knowledge can influence how the perceptual system deals with variance and bias. In such cases, the weight that prior knowledge plays in a perceptual system's determination of maximal invariance is what we have called the perceptual frame. Hence, the perceptual frame determines whether and how multisensory perceptual objects are individuated within a perceiver's sensory environment. Insofar as prior knowledge can shift the probability of veridicality towards familiar forms of organization, perceptual frames modulate what will count as optimal representations of the sensory environment. Consequently, for perceivers who have bias and learned different strategies to address their bias, their perceptual frames may be so bespoke as to yield the perception of very different perceptual objects even given the same set of sensory information.

ii. Amodal completion and mental imagery

The purpose of this section of the paper has been to demonstrate that perceptual frames are a necessary component of perceptual object individuation and, further, how different perceptual frames can yield different perceptual objects given the same sensory input. Above, we have shown how the perceptual system employs principles of optimality to organize multisensory environments. These principles, insofar as they are learned and implemented to address unique cases of perceptual bias, are one example of perceptual framing. In what follows, we shift gears from multisensory perceptual organization to amodal completion, mental imagery, and predictive anticipation in order to demonstrate that perceptual framing also plays a necessary role in the possibility of perceptual object individuation.

Amodal completion is the early perceptual processing and representation of an object without requiring sensory stimulation of some of that object's parts (Nanay, 2018; Thielen et al., 2019). It can occur under a number of different circumstances, but for the purposes of the discussion here we shall focus on this simple definition. The first example of amodal completion included below is a circle occluded by a square. Although the circle appears behind the square and we do not perceive 25% of it, by amodal completion, we nevertheless represent that there *is* a circle behind a square. The second example below is known as the Kanizsa triangle which is an illusion presenting the completion of an upside-down triangle within the negative space of an incomplete, right-side-up triangle. In both cases, the perceptual system completes features that are not present insofar as we represent the complete circle or triangles in the absence of sensory stimulation. Moreover, we may differentiate top-down amodal completion (e.g., object knowledge and expectations shape the completion) from bottom-up amodal completion (e.g., completion of shapes on the basis of Gestalt principles). For our purposes in illuminating the role of perceptual frames in object perception, we are interested in the top-down influenced sort of amodal completion.

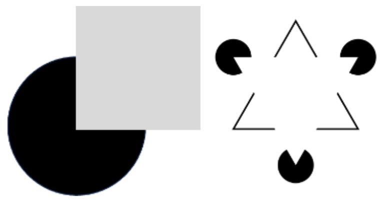

Fig 3. These are two examples of shape completion.[8]

What is important about both of the examples in Figure 3 is that the perceptual system has the ability to represent features despite a lack of sensory stimulation. Further, these representations are distinctly perceptual (as opposed to cognitive) insofar as amodal completion is performed in the early stages of perceptual processing through the primary visual cortex (Ban et al., 2013; Bushnell et al., 2011; Emmanouil and Ro, 2014; Hazenberg, et al., 2014; Lee et al., 2012; Pan et al., 2012; Scherzer and Ekroll, 2015). Initial brain responses to amodal completion emerge 140ms after stimulus presentation (Guttman and Kellman, 2004), followed by differential responses after 240ms (Murray et al., 2004). In both cases, despite fragmentary inputs, we have shape completion processed in the primary visual cortex (V1), which can be explained by intermediate representations of contour interpolation (Kellman and Fuchser, 2023). Importantly, both of the foregoing examples of amodal completion are ubiquitous among perceivers with normally functioning perceptual capacities. However, amodal completion can occur differently for different perceivers when sensory scenes either grow increasingly complex or when there is a dearth of sensory information. Further, these differences can be manipulated in empirical settings demonstrating the effects of perceptual frames on amodal completion.

This shall serve as the basis of our argument as follows: (1) when amodal completion occurs, the 'missing pieces' of a completed object are represented. However, (2) if there were no reason to represent the parts of the object that are missing, then those missing pieces would not be

[8] https://en.wikipedia.org/wiki/Illusory_contours

perceived. So, for example, if you are familiar with what guitars look like and you see a part of one behind an occluding surface, you may amodally complete a representation of the guitar's parts which are not visible to you. However, if you had never seen a guitar before, then there would be no principled basis by which you or your perceptual system might be able to represent whatever mysteries lie on the other side of the occlusion. Accordingly (3), by contraposition, missing pieces of perceptual objects are perceived when there is a reason to represent them. Therefore, if amodal completion occurs, then there is a reason to represent the missing pieces of occluded or non-fully presented perceptual objects. What, then, stands as the 'reason' that explains the representation of amodally completed objects?

Bence Nanay (2010; 2018; 2023) has argued that mental imagery is the foregoing reason for the representation of amodally completed objects. According to Nanay, the imposition of a mental image on a scene completes what is occluded by virtue of how we expect the rest of the occluded object to appear (in the case of vision, at least). As in the example of a partially occluded guitar, knowing what a guitar looks like (having the mental image of a guitar) enables our perceptual system to amodally complete where the object is occluded. Yet, could perceptual frames also help to explain the possibility of amodal completion? We have argued that perceptual frames are distinct from mental imagery insofar as frames are adaptations of the perceptual system that prefer to represent and organize sensory information in particular ways. Whereas mental imagery involves a particular instantiation of perceptual representation not triggered directly by sensory input (Pearson et al. 2015; Dijkstra et al. 2019), the effects of framing are matters of the perceptual system's built-in preferences to appropriate sensory bias and learning for the sake of the veridical organization of sensory data.

Along similar lines, Briscoe (2011) has argued that amodal completion is not a monolithic perceptual phenomenon for which mental imagery is universally sufficient. As he demonstrates, there are cases in which amodal completion can be influenced by top-down processes, yet occurs in a bottom-up manner similar to the effects of perceptual framing – and, thus, without the need for mental imagery. However, Nanay (2023) contends with Briscoe that the application of mental imagery can be unconscious and bottom-up. His evidence comes from two cases – both of which we shall show are better explained by perceptual framing than by unconscious mental imagery. We then argue that where mental imagery is necessary for completion, it relies on the effects of perceptual frames to represent veridically. We conclude alongside Briscoe that not all cases of completion depend upon imagery and further posit that some of those instances are best explained by perceptual frames. We also conclude alongside Nanay that mental imagery is a powerful tool in completion made all the more useful because of the effect that frames have for affording ecologically valid sensory information to processes of mental imagery.

Nanay argues that unconscious mental imagery explains how in cases of aphantasia, a subject can be unable to mentally imagine a shape, but nevertheless be able to discriminate whether random targets on a coordinate grid would fall within the boundaries of that shape were it to be visibly present (Jacobs et. al, 2017). Similarly, regardless of whether subjects were asked to imagine or to repress imagining a red apple, Kwok et. al (2019) found that the subjects were equally primed in either condition to perceptually prefer red cues over alternatives. In both cases, Nanay admits that neither set of subjects could have imposed a top-down mental image in order to excite the priming effects that the researchers found. However, because there were consequent priming effects, Nanay concludes that the subjects employed unconscious, bottom-up mental imagery. To the contrary, neither of the foregoing cases that Nanay introduces as evidence for a bottom-up and unconscious application of mental imagery necessitates such a complex and representationally rich architecture.

Just because an aphantasic cannot mentally imagine a shape does not mean that they are unable to recognize shapes when they see them; nor does it mean that their perceptual systems are unable to collect, organize, and categorize sensory information as optimal, veridically represented objects. Underlying structural principles of perceptual objecthood guide how the perceptual system integrates and organizes sensory information into objects like shapes. Accordingly, the structural principles that organize both normally functioning and aphantasic perceptual experience need not rely on the rich content of imagery to function. Consequently, a perceptual frame may be set such that the aphantasic's perceptual system is predictively readied to organize sensory information in order to represent the shape's distinctive features if they were to appear. Given this framed readiness, targets that appear within the boundaries of the would-be shape do not require a mental image of the shape to be categorized as consistent or inconsistent with it. A similar case can be made for the 'imagine a red apple' study. Rather than an unconscious mental image of an apple causing the measured effects, the perceptual system itself may alternatively be framed to seek and represent individual features associated with red apples. Thus, red cues are preferred not because they are congruent with the mental image of an apple, but rather because the perceptual mechanisms that prefer redness and roundness and shininess are set to high alert by the associated perceptual frame. The mental image of an apple – unconscious or otherwise – is perhaps sufficient, but not necessary for the perceptual system to seek out and organizationally prefer red, round, and shiny properties in the sensory environment.

For this reason, the invocation of mental imagery to explain priming effects and the amodal completion of perceptual objects is an important tool: where mental imagery is utilized, completion can occur. Yet, not all instances of completion require rich mental imagery to operate. In such cases, how the perceptual system is framed can explain why features are primed and organized as they are. Moreover, the framing of the perceptual system to prime and organize sensory environments seems an important pre-requisite of mental imagery in amodal completion. One would not have the opportunity to appropriately apply a mental image were the requisite sensory information ignored by the perceptual system. Mental imagery aids in veridical perception only when images fit the scenes that they aim to complete. Barring hallucination, one cannot veridically apply their mental imagery of northern hemisphere constellations to the southern hemisphere's night sky. So, it would seem that mental imagery is an incomplete explanation for amodal completion. Where it is necessary, one must not only have a mental image available to apply to a scene, but the scene itself must present ecologically valid sensory information for that image to fit. This is why the effects of perceptual frames may serve as an elementary component of amodal completion in addition to mental imagery. Where mental imagery is necessary for amodal completion, perceptual frames are necessary for imagery. Frames guide the pre-requisite functions of the perceptual system to organize sensory information in order that mental images may optimally apply to occluded portions of perceptual objects.

The foregoing arguments implicate that whenever amodal completion occurs, there is a perceptual frame at work enabling non-occluded sensory information to be organized such that the occluded sensory information may be completed with or without a mental image. Perceptual frames shape amodal completion at the level of perceptual processing by shaping which perceptual objects may emerge from the sensory environment by way of mental imagery or otherwise. For instance, a poor-quality x-ray image with missing or blurred details can be differently completed and represented by a novice, a visual artist, an architect, a dentist, or a radiologist. Different domains generate different frames which, in turn, generate unique representations. Thus, the perceptual objects that experts perceive as a result of direct sensory stimulation can be uniquely enriched by amodal completion. Per the foregoing argument, this is because the perceptual frame operates as the principle or reason for which amodal completion occurs.

And you do not need to be a perceptual expert to experience the effects of perceptual framing on amodal completion! One rather provocative example of this is included here as a pair of bimodal images presented in Figures 4 and 5. If you have not seen these images before, you will likely be able to experience the effect of perceptual framing on your ability to amodally complete it. Have a look at Figure 4 first without looking at 5. The image is likely ambiguous and will not mean much to you beyond splotches of black and white. However, once you have done this, look at 5 then back to 4. When you have seen the image in 5, 4 ceases to be ambiguous – you know what you are looking at because you will have acquired the necessary perceptual frame.

Fig 4. A bimodal image. Take a moment to gaze at this picture before moving on. Can you tell what it is? From Teufel et al. (2018), used under Creative Commons CC-BY license.

These images and others like them were used by Teufel et al. (2018) to demonstrate that prior object knowledge can influence how the perceptual system organizes related sensory information. The effect of the images in figures 4 and 5 occurs because after you have seen the full-blown image, “early feature detectors are shaped by high-level object representations. Top-down influences optimize early-information processing mechanisms [within] the current perceptual context” (*ibid:* pg. 6) Or, in other words, your familiarity with the object sets a perceptual frame within vision’s early-processing mechanisms. The frame then operates such that when the relevant contextual visual information is presented, it is organized by virtue of that frame's content. This is only possible after the perceptual frame has been established. Additionally, you cannot unsee what has already been framed and set. The visual confusion you may have experienced when looking at 4 before seeing 5 is no longer accessible once the frame has been settled. So, although you may be applying a mental image to complete the bimodal scene in Figure 4, its perceptual features are organized by virtue of a perceptual frame aimed at presenting the incomplete visual information with as little bias as possible.

Fig 5. The colored in and completed representation of the bimodal image presented in Figure 4. From Teufel et al. (2018), used under Creative Commons CC-BY license.

iii. Predictive processing and anticipation

Predictive processing (PP), may be well suited to explain discussed individual differences in object perception, including varied abilities in sensory integration and perceptual completion in both unimodal and multimodal settings (*for an extensive introduction to PP see* Clark 2015; Hohwy 2013; Metzinger & Wiese 2017). Our perception of the world depends quite substantially upon prediction. According to this explanatory framework, we perceive when our top-down predictions match incoming sensory stimuli. This means that computational processes in the brain importantly consist of interactions between top-down and bottom-up processing underlying cognition, perception, and action. The relative weighting of top-down and bottom-up signals emphasized by this approach is often described as 'hypothesis testing' in the brain, or accumulating evidence for the 'internal model' (Engel et al. 2001; Friston 2005; Clark 2014). The framework operates with sub-personal, computational or neurobiological descriptions but can also be used to analyze personal level descriptions relating to phenomenological reports on the contents of perception, reasoning, and other agentive phenomena (Wiese & Metzinger 2017). Moreover, it emphasizes that the brain is not a passive, stimulus-driven device. Rather our predictive perception of objects is a matter of an active, selective, and constructive process (Wacongne et al. 2011; Bastos et al. 2012; Brodski et al. 2015). This makes PP and its findings a rich ground for evidence of perceptual frames.

Neuroscientific evidence has suggested that the processing of sensory stimuli may importantly be shaped by top-down influences modifying the activity in thalamocortical networks and triggering further predictions about approaching sensory signals (Engel et al., 2001). Higher levels of the cortical hierarchy (e.g., visual word form area VWFA, or fusiform face area FFA) learn not only to track specific sensory information but also to expect sensory features and predict sensory consequences of perceived scenes and events (Reich et al., 2011; Egner et al., 2010). These top-down predictions can be generated in an information-integrating, metamodal manner (Pascual-Leone and Hamilton, 2001). Accordingly, the sensory cortex uses feedback from higher levels of perceptual processing to predict upcoming sensory stimulation (previously transduced by relevant sensory receptors at the lower levels of the processing hierarchy). These sensory processing feedback loops indicate a hierarchical predictive coding model insofar as expectation-induced multimodal context effects are propagated all the way down the perceptual system, priming and changing sensory responses at the lower, so called 'early' processing stages as V1 (Muckli, 2010). The groundbreaking contribution of predictive processing is that it emphasizes the influence of prior knowledge and top-down processing as a predominant and widespread feature of perception (Aitchison and Lengyel, 2017).

Significantly, these top-down effects that modulate early stages of perceptual processing appear to be hardcoded into the perceptual system itself. Empirical studies within predictive processing models of perception have shown that perceptual context, including perceivers' expectations based on previous knowledge about the world, and familiarity with certain perceptual objects, may dynamically interact with the earliest stages of cortical visual processing. For example, these interactions include changes to early feature-detectors in the retina, shaping and sharpening their properties (Teufel et al., 2018; Hosoya et al. 2005). Although frames are typically adaptations of the perceptual system caused by perceptual learning that change emerging object representations in a specific subject-dependent way, it seems reasonable to acknowledge that in some cases these adaptations may also be physically hardwired and evident at the structural level through certain neuronal specializations. Examples of this are special forms of object perception as exhibited by face perception or the phenomenon of synesthesia (Egner et al. 2010; Rouw & Scholte, 2010; Mroczko-Wąsowicz & Nikolić 2013, 2014; Reeder, Sala, & van Leeuwen, 2024).

Further evidence of perceptual framing effects on amodal completion comes from research on anticipation. de Vignemont (2023) compellingly argues that the anticipation of looming objects is a kind of amodal completion for motion. Her argument hinges upon the fact that the primary visual cortex includes a neural mechanism called the 'looming detector' (Regan and Beverley 1978). The function of the looming detector "fixes the direction of an object's motion in 3D space and is preferentially sensitive to increasing or decreasing object size" (de Vignemont 2023: pg. 141). She then argues that these looming detectors operate in early-vision to represent motion and are able to amodally complete the anticipated motion of looming objects. For example, in anticipation of an object flying at you, your looming detectors will operate to amodally complete the object's trajectory as a kind of warning for the unfortunate consequence of not moving yourself away before it strikes.

While cases of looming detection like the foregoing example indicate a hardwired component of normally functioning perceptual capacities, expertise and familiarity appear to be able to perceptually frame looming and anticipation in more complex cases of perceiving object's motion. We can find empirical support for this claim from studies that report significant differences in the visual search behavior and accuracy of anticipatory judgments made by expert athletes compared to novices (Abernethy and Koning, 1994; Abernethy et al., 2012; Murphy et al., 2016; Runswick et al., 2018; Thomas et al., 2022). Experts are not only better at amodally completing future-anticipated events, but they also rely on different sets of visual information than non-experts even when being afforded the same sensory stimulation.

What these data demonstrate is that experts have a superior capacity to anticipate future outcomes within their domains of expertise than novices. All of this taken together suggests that, if anticipation is a kind of amodal completion, then the ability to amodally complete an anticipated event depends upon familiarity and experience. Expertise, therefore, involves the implementation of perceptual frames that provide the ability to amodally complete in anticipation of future events. This remains in accordance with the following insights of the predictive coding framework.

Namely, expectations speed up reportable sensory awareness of perceived objects or events; and, so predicted stimuli are correlated with enhanced visibility or more generally with an increased performance (Melloni et al. 2011; cf. Muckli 2010). As Andy Clark (2014:36) puts it: "(…) the familiar song *really does* sound clearer. It is not that memory *later* does some filling-in that affects, in some backward-looking way, how we judge the song to have sounded. Rather, the top-down effects bite in the very earliest stages of processing, leaving us little conceptual space (…) to depict the effects as anything other than enhanced-but-genuine perception." Therefore, one's accurate perception in cases of predictive amodal completion depends upon having a perceptual frame to guide the reception and organization of afforded sensory information.

Just as in Figure 4 without the proper context (or perceptual frame), the perceptual system is unable to organize a meaningful representation of the bimodal image. Or, at least, it is unable to represent the bimodal image as anything other than an abstract set of black and white splotches. The transformation of those splotches into a coherent image depends crucially upon having the appropriate perceptual frame. Similarly, in predictive amodal completion, without the proper perceptual frame, novices cannot hope to anticipate with the accuracy of experts who know what is to come next on the basis of their readiness to perceive it. Thus, we have demonstrated a difference in the representation of perceptual objects on the basis of having applied a perceptual frame (a specific top-down prediction or prior knowledge) rather than from a change in sensory input. The input to the perceptual system remains the same, but the matter of its framed organization changes the objects that are perceived.

IV. Perceptual frames across the joint of cognition and perception

So far, we have demonstrated how perceptual frames make substantial contributions to the processes of object perception. These demonstrations yielded two important insights. First, perceptual frames can cause differences in multisensory integration that lead to the possibility of different perceptual objects being formed from the same sensory information. Second, perceptual frames underlie the predictive capability of representing perceptual objects in amodal completion. Significant to both of these insights is that they reveal ways in which low-level sensory processes are responsive to prior information stored in the perceptual system. That is, once implemented into the perceptual system (often in a top-down manner), perceptual frames operate alongside other structural principles of perceptual objecthood to organize and represent coherent sensory scenes.

This highlights the third, and perhaps most important, insight that we develop in this final section of the paper: The effects of perceptual framing implicate that the perceptual system *as a whole* is responsive to adaptation from top-down, cognitive influence; and yet, the perceptual frames themselves appear to operate in a largely bottom-up fashion.

What we learn, do, and encounter in our lives impacts what we can perceive. The aspects of a sensory scene that we perceptually access depend upon our efforts and goals. Thus, cognitively-driven perceptual frames remind us about perceptual diversity. The investigations of perceptual framing effects suggest that our perceptual experiences of objects and related object-directed perceptual capacities cannot be considered universal. The contents of object representations differ across subjects and times. This is, however, not to say that our perceptual experiences depend constitutively on our cognitive states, nor that we perceive what we think. This is why our perception is compatible with intersubjective objectivity and the mind-independence of the objects we perceive (Longino, 1990; O'Callaghan, 2024). Cognition can drive diversity in perceptual processing while that processing remains bottom-up.

The relationship between top-down and bottom-up effects as well as the potential overlap of cognitive operations with perceptual processing remain widely contentious and diverse subjects of debate (Shea, 2015). To briefly summarize, it is assumed that the mind and its processes exhibit a "joint" in nature between perceptual states and cognitive states such as thoughts, intentions, beliefs, and desires (Firestone and Scholl, 2016; Block, 2023). Whether the joint between perception and cognition is permeable is the big point of disagreement. Some argue that perception remains unaffected by cognition through 'encapsulation' or 'modularity' (Fodor, 1983; Orlandi, 2014; Firestone and Scholl, 2016), while others argue that it is possible for cognition to penetrate perception by affecting either its processes or the contents of perceptual experience (Bhalla and Proffit, 1999; Stefanucci and Guess, 2009). As we mentioned at the beginning of this paper, we remain neutral with respect to this debate. We have nothing to say here about whether individual instances of perception are penetrable by cognition. However, the third insight from perceptual frames mentioned above does offer an important point of inflection for this debate about encapsulation and permeability. Namely, perceptual frames invite us to think about the perceptual system as a whole apart from individual instances of perception.

For the strong modularists like Firestone and Scholl (2016: pg. 17), it is common to find claims similar to the following: “There is a joint between perception and cognition and the nature of this joint is such that perception proceeds without any direct, unmediated influence from cognition.” Our discussion of perceptual frames adds a new perspective to a growing sentiment in the literature that this kind of claim is inaccurate. That is, acknowledging the border between perception and cognition does not have to lead to questioning cognitive penetration (Green, 2020; Block, 2023). In fact, cognition does have a direct and unmediated influence on perception insofar as it is through top-down influences that perceptual frames are encoded into the perceptual system. As we learned from the discussion of MLE, the organizational principles of low-level perceptual processes are sensitive to prior knowledge in optimally representing multisensory environments. Similarly, we learn from amodal completion that perceptual frames rely on specific object-knowledge to determine which perceptual objects are possible to be represented at all. So – perceptual frames show that there is a direct influence of cognition on perceptual processing which adapts over time and with experience. And yet, even if the foregoing modularist claim is false, this does not mean that modularists are wrong about their views regarding impenetrability, nor contrastingly that cognitive penetration is real. Rather, the case from perceptual frames invites nuance into how we ought to consider the joint between perception and cognition (Mroczko-Wąsowicz, 2023).

What we learn from perceptual frames is that the perceptual system *as a whole* is adaptive and receptive to cognitive influence. It is through cognitive influence that perceptual frames are diachronically invested into the perceptual system. Perceptual frames are evidence of adaptations in perceptual processing and the data from multisensory integration and amodal completion indicate that top-down predictions are a driving force that cause those adaptations to occur. Yet, the cognitive origin that explains why a perceptual frame was implemented in the first place need not be the same reason why that frame operates in early sensory processing. Nothing about the functioning of a perceptual frame in any *individual instance* of perception necessitates the same kind of etiology that caused the frame to be installed originally. After all, the object knowledge that facilitates creating a perceptual frame is not necessarily what causes the perceptual system to prefer that object in a sensory environment. This can be explained by framed predictive processing on its own.

The effects of perceptual frames may well be as bottom-up and encapsulated from cognition as any modularist would accept (*see* Mroczko-Wąsowicz, 2022 *for a review*). And, in keeping with our neutrality, nothing about perceptual frames necessitates that they be impenetrable by cognition. Rather, the big takeaway is that there are different levels of description with reference to the joint between cognition and perception. What we have demonstrated is that at a holistic, system-wide level of description, cognition does affect perception by implementing perceptual frames. In contrast, whatever we might say about the perceptual system as a whole, whether individual episodes of perception are penetrable by cognition remains an open question because the operations of perceptual frames are bottom-up.

To wrap up these discussions, we may return to our example scenario and the corresponding question with which we introduced this paper: "When listening to the concerto and perceiving the BACH motif, does the professor of music hear anything different than you?" Without invoking the suppositions of either modularity or penetrability, it is possible to answer "yes" to the foregoing question by reference to perceptual frames. In such a case, the perceptual expert will have a perceptual frame that attunes their perceptual system to prefer the arrangement of the notes that constitute the BACH motif. They will likely also perceive many of the same perceptual objects as other perceivers, but the expert's perceptual frame makes them sensitive to an object of perception (the motif) that others without the appropriate frame could not perceptually individuate. While it may be post-perceptually possible to re-arrange one's experience of the music to understand what the BACH motif sounded like, the expert's automatic reception of it demonstrates a difference in object perception without a difference to the inputs of perception. This, in turn, is an exemplification of the fact that the perceptual system is adaptive. The mosaic of possible objects for perception in any given scene depends upon who is perceiving it and how tailored their perceptual system has become.

V. Conclusion

What we have demonstrated in this paper has significance for the debate concerning the influence of cognition on perception across their "joint." The influence of perceptual frames on the perceptual system *as a whole* is genuinely cognitive. It takes an expert to see or otherwise experience perceptual objects within their domain as only experts can. It takes top-down guidance and the appropriate context to organize bimodal smudges into a coherent image of a woman and a horse (for example). However, what makes a perceptual frame a unique kind of cognitive influence on perception is that it is an encoded mechanism implemented into the perceptual system such that it operates in a bottom-up fashion. Perceptual frames, in this manner of speaking, are something like perceiver-specific structural principles of perception.

For the foregoing reasons, cognition and perception may remain encapsulated from one another while also being deeply intertwined because the perceptual system is adaptive. Its adaptations are perceptual frames, and the principle by which it adapts is, in part, inspired by cognition. Consequently, even if individual instances of perception remain unaffected by cognition, the perceptual system, on the whole, can be. That is the power of the perceptual frame. This invites further nuance for how we discuss the joint between cognition and perception. Although the perceptual system as a whole is adaptive and therefore responsive to cognitive influence, it remains an open question whether the same is true of individual perceptual experiences. We hope that this introduction of perceptual frames will help to better understand the joint across which perception and cognition communicate to represent the perceptual world.

Acknowledgments

This work was funded by the National Science Centre, Poland (AMW's grant: 2019/35/B/HS1/04386). For the purpose of Open Access, the authors have applied for a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. Additionally, we would like to thank our anonymous reviewers who graciously helped improve the quality of this paper through the revision process.

References

Abernethy, B., Neal, R.J. and Koning, P. (1994), Visual–perceptual and cognitive differences between expert, intermediate, and novice snooker players. *Appl. Cognit. Psychol.*, 8: 185-211. https://doi.org/10.1002/acp.2350080302

Aitchison L, Lengyel M. (2017). With or without you: predictive coding and Bayesian inference in the brain. *Curr Opin Neurobiol*. 46:219-227. doi: 10.1016/j.conb.2017.08.010.

Ban, H., Yamamoto, H., Hanakawa, T., Urayama, S., Aso, T., Fukuyama, H., & Ejima, Y. (2013). Topographic representation of an occluded object and the effects of spatiotemporal context in human early visual areas. *Journal of Neuroscience*, 33, 16992–17007. doi: 10.1523/JNEUROSCI.1455-12.2013.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P. & Friston, K. J. (2012). Canonical microcircuits for predictive coding. Neuron, 76 (4), 695-711. https:// dx.doi.org/10.1016/j.neuron.2012.10.038.

Bhalla, M. & Proffitt, D. R. (1999) Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance* 25:1076–96. doi: 10.1037//0096-1523.25.4.1076.

Bermejo, F., Hüg, M. X., & Di Paolo, E. A. (2020). Rediscovering Richard held: Activity and passivity in Perceptual Learning. *Frontiers in Psychology, 11*. doi: 10.3389/fpsyg.2020.00844

Block, N. (2010). Attention and Mental Paint. *Philosophical Issues*, 20, 23–63.

Block, N. (2023). *The border between seeing and thinking*. Oxford University Press.

Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A.M., Helsen, W.F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. *Psychol Bull*. 145(10):980-1027. doi: 10.1037/bul0000207.

Briscoe, R. (2011). Mental imagery and the varieties of amodal perception. *Pacific Philosophical Quarterly*. 92 (2):153-173.

Brodski, A., Paasch, G.-F., Helbling, S. & Wibral, M. (2015). T he faces of predictive coding. The Journal of Neuros cience, 35 (24), 8997-9006. https://dx.doi.org/10.1523/ jneurosci.1529-14.2015.

Bushnell, B. N., Harding, P. J., Kosai, Y., & Pasupathy, A. (2011). Partial occlusion modulates contour-based shape encoding in primate area V4. *Journal of Neuroscience*, 31, 4012–4024. doi: 10.1523/JNEUROSCI.4766-10.2011.

Carrasco, M., Giordano, A.M., and McElree, B. (2004). Temporal performance fields: visual and attentional factors. *Vision Research,* 44(12): 1351–1365. https://doi.org/10.1016/j.visres.2003.11.026

Carrasco, M., Giordano, A.M., and McElree, B. (2006). Attention speeds processing across eccentricity: feature and conjunction searches. *Vision Research*, 46(13): 2028–2040. https://doi.org/10.1016/j.visres.2005.12.015

Carrasco, M. (2018). How visual spatial attention alters perception. *Cogn Process*., 19(Suppl 1):77-88. doi: 10.1007/s10339-018-0883-4.

Casati, R., (2015). Object perception. In Mohan Matthen (ed.), *The Oxford Handbook of the Philosophy of Perception*. Oxford University Press UK.

Chalmers, D. (2004). The Representational Character of Experience. In B. Leiter (Ed.), *The Future for Philosophy*. Oxford: Oxford University Press.

Clark, A. (2000). *A Theory of Sentience*. New York, NY: Oxford UP.

Clark, A. (2014). Perceiving as Predicting. In D. Stokes, M. Matthen & S. Biggs (Eds.) Perception and Its Modalities. Oxford University Press, New York, pp. 23-43 .

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford University Press, UK.

Cohen, J. (2023) Multimodal binding as mereological co-constituency. In A. Mroczko-Wąsowicz, and R. Grush (Eds.), *Sensory Individuals: Unimodal and Multimodal Perspectives*. Oxford, UK: Oxford University Press.

de Vignemont, F. (2023). Looming Perception: Seeing in a Dynamic World. In: A. Mroczko-Wąsowicz and R. Grush (Eds.) *Sensory Individuals: Unimodal and Multimodal Perspectives*. Oxford: Oxford University Press.

Dijkstra, N., S. E. Boschand M. A. J. van Gerven, (2019). Shared neural mechanisms of visual perception and imagery. *Trends in Cognitive Sciences*, 23: 423–434.

Drew, T., Evans, K., L-H Võ, M., Jacobson, F., and J Wolfe, J. (2013) Informatics in Radiology: What Can You See in a Single Glance and How Might This Guide Visual Search in Medical Images? *RadioGraphics,* 33(1): 263-274. doi: 10.1148/rg.331125023.

Egner, T., Monti, J. M., Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. Journal of Neuroscience, 30(49): 16601-16608.

Emmanouil, T., & Ro, T. (2014). Amodal completion of unconsciously presented objects. *Psychonomic Bulletin & Review*, 21, 1188–1194. doi: 10.3758/s13423-014-0590-9

Engel, A. K., Fries, P. & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. Nat Rev Neurosci, 2 (10), 704–716.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature,* 415, 429–433. https://doi.org/10.1038/415429a

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences,* 8, 162–169. https://doi.org/10.1016/j.tics.2004.02.002

Ernst, M. O. (2012). Optimal multisensory integration: Assumptions and limits. In B. E. Stein (Ed.), *The new handbook of multisensory perception*. Cambridge, MA: MIT Press. https://doi.org/10.1016/j.tics.2004.02.002

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral & Brain Sciences*, *39:e229,* 1-77. doi: 10.1017/S0140525X15000965

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society B: Biological Scien ces, 360 (1456), 815-836. https://dx.doi.org/10.1098/rstb.2005.1622.

Gegenfurtner, A., Lehtinen, E., and Säljö, R. (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educ. Psychol. Rev*., 23(4)523–552. https://doi.org/10.1007/s10648-011-9174-7

Green, E. J. (2018) A theory of perceptual objects. *Philosophy and Phenomenological Research*, 99(3),663–93. https://doi.org/10.1111/phpr.12521

Green, E. J. (2020). The Perception-Cognition Border: A Case for Architectural Division. *Philosophical Review* 129 (3):323-393. https://doi.org/10.1215/00318108-8311221

Green, E.J., & Schellenberg, S. (2018). Spatial perception: The perspectival aspect of perception. *Philosophy Compass, 13:e12472*. https://doi.org/10.1111/phc3.12472

Guttman, S. E. & Kellman, P. J. (2004). Contour interpolation revealed by a dot localization paradigm. *Vision Research*, 44(15), 1799– 815. https://doi.org/10.1016/j.visres.2004.02.008

Haider, H. & Frensch, P. A. (1999). Eye movement during skill acquisition: more evidence for the information-reduction hypothesis. *J. Exp. Psychol.: Learn. Memory,* 25, 172–190. https://doi.org/10.1037/0278-7393.25.1.172

Hazenberg, S. J., Jongsma, M. L., Koning, A., & van Lier, R. (2014). Differential familiarity effects in amodal completion: Support from behavioral and electrophysiological measurements. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 669–684. doi: 10.1037/a0034689

Helbig, H. B., & Ernst,M. O. (2007). Knowledge about a common source can promote visual haptic integration. *Perception*, 36, 1523–1533. https://doi.org/10.1068/p5851

Held, R. and Bossom, J., (1961). Neonatal Deprivation and Adult Rearrangement: Complementary Techniques for Analyzing Plastic Sensory-Motor Coordinations. *Journal of Comparative and Physiological Psychology*, 54(1): 33–37. https://doi.org/10.1037/h0046207

Held, R., (1961). Exposure-History as a Factor in Maintaining Stability of Perception and Coordination. *The Journal of Nervous and Mental Disease*, 132(1): 26–32. https://doi.org/10.1097/00005053-196113210-00005

Helmholtz, H, (2005) [1924]. *Treatise on Physiological Optics,* (Volume 3). New York: Dover.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hosoya, T., Baccus, S.A., Meister, M. (2005). Dynamic predictive coding by the retina. *Nature* 7;436(7047):71-7. doi: 10.1038/nature03689.

Ivy, S. (2023). Unconscious Intelligence in the Skilled Control of Expert Action. *Journal of Consciousness Studies* 30 (3):59-83.

Ivy, S., T. Rohovit, M. Lavelle, L. Padilla, J. Stefanucci, D. Stokes, and T. Drew. (2021). Through the eyes of the expert: Evaluating holistic processing in architects through gaze-contingent viewing. *Psychonomic Bulletin & Review,* 28: 870-878. https://doi.org/10.3758/s13423-020-01858-w

Ivy, S., Rohovit, T., Stefanucci, J., Stokes, D., Mills, J., & Drew, T. (2023). Visual expertise is more than meets the eye: an examination of holistic visual processing in radiologists and architects. *J. Med. Imag*. doi: 10.1117/1.JMI.10.1.015501

Jackson, F. (1977). *Perception: A Representative Theory*. Cambridge: Cambridge UP.

Jacobs, C., D. S. Schwarzkopf, and J. Silvano. (2017). Visual Working Memory Performance in Aphantasia*. Cortex 105: 61–73.*

Kellman, P.J., Fuchser, V. (2023). Visual completion and intermediate representations in object formation. In A. Mroczko-Wąsowicz, and R. Grush (Eds.), *Sensory Individuals: Unimodal and Multimodal Perspectives*. Oxford, UK: Oxford University Press.

Kirsch, W., & Kunde, W. (2023). On the Role of Interoception in Body and Object Perception: A Multisensory-Integration Account. *Perspectives on Psychological Science*, *18*(2), 321-339. doi: 10.1177/17456916221096138

Kundel, H., Nodine, C., Conant, E., Weinstein, S. (2007). Holistic Component of Image Perception in Mammogram Interpretation: Gaze-Tracking Study. *Radiology* 242(2): 396-402. doi: 10.1148/radiol.2422051997

Kwok, E. L., G. Leys, R. Koenig-Robert, and J. Pearson. (2019). Measuring Thought-Control Failure: Sensory Mechanisms and Individual Differences. *Psychological Science.* 30: 811–21.

Longino, Helen E. (1990). Science as Social Knowledge: Values and Objectivity in Scientific Inquiry. Princeton University Press.

Mandrigin, A. (2018). Multisensory Integration and Sense Modalism. *The British Journal for the Philosophy of Science,* 2021 72:1, 27-49. https://doi.org/10.1093/bjps/axy070

Matthen, M. (2023). Material objects as the singular subjects of multimodal perception. In A. Mroczko-Wąsowicz and R. Grush (eds.) *Sensory Individuals: Unimodal and Multimodal Perspectives,* Oxford, UK: Oxford University Press.

Meng, M., Tong F, Blake R. (2006). Neural bases of binocular rivalry. *Trends Cogn Sci.*, 10(11):502-11. doi: 10.1016/j.tics.2006.09.003

Metzinger, T., Wiese, W. (Eds.). (2017). Philosophy and Predictive Processing: 1. Frankfurt am Main: MIND Group. doi: 10.15502/9783958573024

Mroczko-Wąsowicz, A., Nikolić, D. (2013). Colored alphabets in bilingual synesthetes. In J. Simner & E. Hubbard (Eds.), *Oxford Handbook of Synesthesia* (pp. 165–180). Oxford: Oxford University Press.

Mroczko-Wąsowicz, A., Nikolić, D. (2014). Semantic mechanisms may be responsible for developing synesthesia. *Frontiers in Human Neuroscience,* 8:509. doi: 10.3389/fnhum.2014.00509

Mroczko-Wąsowicz, A. (2022). Modularity. In B. D. Young & C. Dicey Jennings (Eds.). *Mind, Cognition, and Neuroscience: A Philosophical Introduction* (pp. 149-163). New York: Routledge Press. ISBN 9781138392366

Mroczko-Wąsowicz, A. (2023). Perceptual expertise and object recognition: An explanatory task for modularists and antimodularists. *Philosophy and the Mind Sciences, 4, 14. https://doi.org/10.33735/phimisci.2023.10247*

Mroczko-Wąsowicz, A., Grush, R. (2023). Introduction: Sensory individuals: Contemporary perspectives on modality-specific and multimodal objecthood. In: A. Mroczko-Wąsowicz and R. Grush (Eds.) *Sensory Individuals: Unimodal and Multimodal Perspectives* (pp.1-16). Oxford: Oxford University Press.

Mroczko-Wąsowicz, A., Stoch, N., Zguda, P. (2023). What makes something a perceptual object? In: A. Mroczko-Wąsowicz and R. Grush (Eds.) *Sensory Individuals: Unimodal and Multimodal Perspectives* (pp. 37–54). Oxford: Oxford University Press.

Mroczko-Wąsowicz, A., Ivy, S., Bachanek, M., & Cząstkiewicz, A. (2024). Accounting for Action: Challenging the Traditional View of Multimodal Perceptual Objects. In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the Annual Meeting of the 46th Annual Conference of the Cognitive Science Society.*

Muckli, L. (2010) What Are We Missing Here? Brain Imaging Evidence for Higher Cognitive Functions in Primary Visual Cortex V1 IJIST 20: 131-139.

Murray M. M., Foxe D. M., Javitt D. C., & Foxe J. J. (2004). Setting boundaries: Brain dynamics of modal and amodal illusory shape completion in humans. *Journal of Neuroscience*, 24: 6898–6903. doi: 10.1523/JNEUROSCI.1996-04.2004

Nanay, B. (2010). Perception and Imagination: Amodal Perception as Mental Imagery. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*,150 (2), pp. 239–54.

Nanay, B. (2018). The Importance of Amodal Completion in Everyday Perception. *I-Perception,* 9(4): 1–16. doi: 10.1177/2041669518788887

Nanay, B. (2023). *Mental Imagery: Philosophy, Psychology, Neuroscience*. Oxford University Press.

O'Callaghan, C. (2008). Object Perception: Vision and Audition. *Philosophy Compass,* 3: 803–29. https://doi.org/10.1093/acprof:oso/9780198782964.003.0003

O'Callaghan, C. (2016). Objects for Multisensory Perception. *Philosophical Studies,* 173: 1269–89. https://doi.org/10.1007/s11098-015-0545-7

O'Callaghan, C. (2024). Perceptual expertise, universality, and objectivity. *Philos Stud* 181, 1757–1763. https://doi.org/10.1007/s11098-023-02019-0

Ongchoco, J.. & Scholl, B. (2022). Hallucinating visual structure: Individual differences in 'scaffolded attention'. *Cognition* 225 (C):105129.

Ongchoco, J., & Scholl, B. (2023). Figments of imagination: 'Scaffolded attention' creates non-sensory object and event representations. In A. Mroczko-Wąsowicz, and R. Grush (Eds.), *Sensory Individuals: Unimodal and Multimodal Perspectives*. Oxford, UK: Oxford University Press.

Orlandi, N. (2014). *The Innocent Eye: Why Vision is not a cognitive process*. Oxford University Press.

Pan, Y., Chen, M., Yin, J., An, X., Zhang, X., Lu, Y., Wang, W. (2012). Equivalent representation of real and illusory contours in macaque V4. *The Journal of Neuroscience*, 32, 6760–6770. https://doi.org/10.1523/JNEUROSCI.6140-11.2012

Pascual-Leone, A., and Hamilton, R. (2001). The metamodal organization of the brain. Progress in Brain Research. 134, 427–445.

Pearson, J., T. Naselaris, E. A. Holmes, and S. M. Kosslyn, (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, 19: 590–602.

Reich, L., Szwed, M., Cohen, L., and Amedi, A. (2011) A ventral stream reading center independent of visual experience Current Biology 21, 363-368.

Lee, S. H., Kwan, A. C., Zhang, S., Phoumthipphavong, V., Flannery, J. G., Masmanidis, S. C., Yang, D. (2012). Activation of specific interneurons improves V1 feature selectivity and visual perception. *Nature*, 488, 379–383. https://doi.org/10.1038/nature11312

Reeder, R.R., Sala, G., van Leeuwen, T.M. (2024). A novel model of divergent predictive perception. *Neurosci Conscious.* 2024(1):niae006. doi: 10.1093/nc/niae006.

Rouw, R., Scholte, H.S. (2010). Neural basis of individual differences in synesthetic experiences. *J Neurosci.* 30(18):6205-13. doi: 10.1523/JNEUROSCI.3444-09.2010.

Scherzer, T. R., & Ekroll, V. (2015). Partial modal completion under occlusion: What do modal and amodal percepts represent. *Journal of Vision*, 15, 1–20. doi:https://doi.org/10.1167/15.1.22

Schwenkler, J. (2014). Vision, Self-Location, and the Phenomenology of the 'Point of View'. *Noús,* 48 (1): 137–155. doi: 10.1111/j.14680068.2012.00871.x

Shea, N. (2015). Distinguishing top-down from bottom-up effects. In D. Stokes, M. Matthen, & S. Biggs (Eds.), Perception and its modalities (pp. 73–91). Oxford University Press.

Sheridan, H. and Reingold, E.M., (2017). The holistic processing account of visual expertise in medical image perception: a review. *Front. Psychol*., 8, 1620. doi: 10.3389/fpsyg.2017.01620

Skrzypulec, B (2022). Subject-dependency of perceptual objects. *Erkenntnis* 87:2827–2842. https://doi.org/10.1007/s10670-020-00328-y

Stefanucci, J. K. & Geuss, M. N. (2009) Big people, little world: The body influences size perception. *Perception* 38:1782–95. doi: 10.1068/p6437

Stratton, G.M., (1896). Some Preliminary Experiments on Vision without Inversion of the Retinal Image. *Psychological Review*, 3: 611–617. https://doi.org/10.1037/h0072918

Stratton, G.M., (1897). Vision Without Inversion of the Retinal Image. *Psychological Review*, 4(4): 341–360 and 4(5)463–481. https://doi.org/10.1037/h0071173

Teufel, C., S. C. Dakin, and P. C. Fletcher. (2018). Prior Object-Knowledge Sharpens Properties of Early Visual Feature-Detectors. *Scientific Reports*, 8: 10853. https://doi.org/10.1038/s41598-018-28845-5

Thielen, J., Bosch, S. E., van Leeuwen, T. M., van Gerven, M. A. J., & van Lier, R. (2019). Neuroimaging Findings on Amodal Completion: A Review. *I-Perception*, *10*(2). https://doi.org/10.1177/2041669519840047

Treisman, A. (1998). The perception of features and objects. In R. D. Wright (Ed.), *Visual attention* (pp. 26–54). Oxford University Press.

Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent surfaces. *Vision Research*, 45, 1095–1098. doi: 10.1016/j.visres.2004.11.001

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L. & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci U S A, 108 (51), 20754-9. https://dx.doi.org/10.1073/pnas.1117807108.

Wagemans, J., Elder, J. H., Kubovy, M., et al. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure- ground organization. *Psychological Bulletin*, 138(6), 1172– 217. doi: 10.1037/a0029333

Wiese, W. & Metzinger T. (2017). Vanilla PP for Philosophers: A Primer on Predictive Processing. In T. Metzinger & W. Wiese (Eds.). Philosophy and Predictive Processing: 1. Frankfurt am Main: MIND Group. doi: 10.15502/9783958573024

Wu, Wayne (2014). *Attention*. New York: Routledge.