



# Visual attention in pictorial perception

Gabriele Ferretti<sup>1</sup> · Francesco Marchi<sup>2</sup>

Received: 8 January 2020 / Accepted: 11 September 2020 / Published online: 10 October 2020  
© Springer Nature B.V. 2020

## Abstract

According to the received view in the philosophical literature on pictorial perception, when perceiving an object in a picture, we perceive both the picture's surface and the depicted object, but the surface is only unconsciously represented. Furthermore, it is suggested, such unconscious representation does not need attention. This poses a crucial problem, as empirical research on visual attention shows that there can hardly be any visual representation, conscious or unconscious, without attention. Secondly, according to such a received view, when looking at a picture aesthetically, one both consciously represents and visually attends to both the depicted object and the picture's surface simultaneously. Thus, contra the empirical research on attention, only conscious visual representations are coupled, by such current view, with attention. And this clearly poses a second problem, as this philosophical account is not in tune with what vision science tells us about the functioning of our visual system. Furthermore, this raises another crucial problem, namely, that of explaining why aesthetic experience of pictures does not feel odd or conflicting, since, as previously noted in the philosophical literature, and contra the received view, if we are simultaneously consciously perceiving both the picture's surface and the depicted object, there seems to be two things, at the same time, in the foreground of one's visual consciousness. But, if so, as suggested, this would lead to a conflicting spatial visual experience. This paper offers a new description of the role of visual attention in picture perception, which explains the difference between the usual and the aesthetic way of perceiving a depicted object, without facing the problems reported above. A crucial role in our new account is played by the notion of unconscious attention, the distinction between focal and distributed, as well as top-down and bottom-up visual attention and the relationship between visual attention and visual consciousness. The paper, thus, offers the first theory concerning the exercise of visual attention in pictorial perception that is both philosophically rigorous and empirically reliable.

---

✉ Gabriele Ferretti  
fairg@live.it

Extended author information available on the last page of the article

**Keywords** Picture perception · Pictorial simultaneity · Visual attention · Picture surface · Depicted object · Vision · Visual consciousness · Aesthetic appreciation of pictures

Perceiving an object in a picture is a common everyday experience. However, there are important phenomenological, neural and behavioural differences between perceiving an object in a picture and perceiving a real object.<sup>1</sup> In this respect, a full understanding of the nature of such a perceptual state still demands a coherent explanation, capable of satisfying both cognitive scientists and philosophers.<sup>2</sup> A fundamental problem in developing such an explanation is that of understanding the role of visual attention in pictorial perception. This topic, however, has been poorly investigated in the literature, with the notable exception of a discussion concerning the exercise of attention in aesthetic appreciation of pictures (Nanay 2016, 2017). In this article, we start from this attempt, and propose a new theory of visual attention in picture perception. Before cashing out our theory, however, we need to briefly review the state of the art about this subject matter.

## 1 Picture perception: the state of the art

Investigating the nature of picture perception is one of the most important challenges for those interested in the study of visual experience.<sup>3</sup> The central and broad question in the literature is the following:

What kind of visual state are we in when we see an object in a picture?

Most philosophers agree that one of the crucial characteristics of picture perception is that we see not one, but two things: the depicted object, i.e. the pictorial content, and the picture's surface, i.e. the vehicle conveying the pictorial content.<sup>4</sup> Now, it seems intuitive that, when the depicted object is, for example, a cherry, we see it just because we are visually representing specific visual properties of the surface in which the depicted cherry is visually encoded: we see the depicted cherry by seeing the marks realized across the surface, e.g. red brush marks in the case of a painting, which convey the pictorial content. Such marks are grouped by our visual representations into a cherry-shaped visual object. In this respect, the way in which the marks on the picture's surface are realized, and consequently visually represented, will shape the way we visually experience the cherry in the pictorial space.<sup>5</sup>

At this point, one might be tempted to conclude that one sees the surface simply because one sees what it encodes, i.e. by seeing the portion of it where the object is depicted. However, such a material dependence of the pictorial content on the

<sup>1</sup> Nanay (2011, 2015a, b, 2017); Lopes (2005), Ferretti (2016a, b, 2017b, 2018a; b, 2019, 2020a, b, Matthen (2005), Hopkins (2003, 2010) and Kulvicki (2006).

<sup>2</sup> See Ferretti (2016b, 2017b, 2018a, 2020a, b).

<sup>3</sup> Kulvicki (2006), Hopkins (2003, 2010), Nanay (2010b, 2011, 2015a, 2017), Newall (2011), Voltolini (2015), Chasid (2014), Aasen (2014), Gregory (2012) and Ferretti (2016b, 2017b, 2018a).

<sup>4</sup> For recent reviews see Nanay (2015a, b, 2017), Ferretti (2018a, b, 2019, 2020a, b).

<sup>5</sup> Voltolini (2013).

pictorial vehicle, despite its intuitive plausibility, does not entail any sort of perceptual dependence: from the notion that one perceives the depicted cherry because it is visually encoded in the surface, whose properties can be cherry-grouped, we cannot infer, *ipso facto*, that one also perceives the surface as such. As Nanay notes: “[...] just because the picture surface is right in front of us, this does not mean that it is perceptually represented. An alternative would be to say that we only represent the depicted object perceptually—the picture surface is not perceived at all” (Nanay 2017: footnote 5; see also 2016: p. 41). Similarly, Lopes contends that: “It is only in virtue of seeing the configuration of marks on its surface (...) that we see anything at all in the picture. However, seeing a pictorial design face to face does not entail seeing the design as a design” (Lopes 2005: p. 28).

All this suggests that one’s being in front of the surface does not entail that she also perceives it as the vehicle that conveys the pictorial content. We need an argument in order to defend the idea that, on top of visually representing the depicted object, we also visually represent the surface.<sup>6</sup> If so, provided that, during picture perception, we visually represent the depicted object, the fundamental question is whether we also always visually represent the surface.

Several philosophical and empirical arguments have been recently offered to suggest a positive answer to this question.<sup>7</sup> One for all, if we were not visually representing the presence of a surface at all, we would enter the pictorial illusion fostered by *trompe l’oeils*, in which the depicted object looks like a present object we can motorically interact with (Ferretti 2018a, 2020a, b). But the idea that we always visually represent the surface is crucially related to the question on whether the visual representation of the depicted object and the visual representation of the surface occur simultaneously, or, instead, our visual system alternates between these two visual states. The literature about this problem mirrors the venerable dispute between Wollheim (1987, 1980, 1998), who defended the idea of *simultaneity*, and Gombrich (1960), who argued for the *alternation* of these visual states.<sup>8</sup> At the moment, most philosophers hold that we *simultaneously* visually represent both the surface and the depicted object.<sup>9</sup>

So far so good. However, things become more complicated when one wants to avoid talking in general terms of ‘seeing’, as well as of the notion of ‘visual representation’.<sup>10</sup> Indeed, seeing can be conscious or *unconscious*<sup>11</sup>—for example in blindsight (Kentridge 2004; Kentridge et al. 2008) or hemispatial visual neglect (Ro and Rafal 1996). At this point, however, one should specify whether one means con-

<sup>6</sup> For a review of this crucial point in the literature, see Ferretti (2018a, 2018b, 2019, 2020a, b).

<sup>7</sup> For a review see Nanay (2011, 2015a, 2017), Ferretti (2018a, b, 2019, 2020a, b).

<sup>8</sup> For a recent review of this debate see Nanay (2011, 2015a, b, 2017), Hopkins (2012), Briscoe (2016, 2018), Ferretti (2018a,b, 2019, 2020a, b).

<sup>9</sup> For a review see Nanay (2011, 2015a, 2017), Ferretti (2018a, b, 2019, 2020a, b).

<sup>10</sup> ‘Seeing’ and ‘visually representing’ can be used interchangeably in this specific literature. For a general and recent review of the different commitments to these notions see (Ferretti and Glenney 2020).

<sup>11</sup> See, Cattaneo and Vecchi (2011), Ferretti and Marchi (2020), Ferretti (2017c), Ferretti and Glenney (2020).

sciously or unconsciously ‘seeing’, or ‘visually representing’, the depicted object and the surface simultaneously.<sup>12</sup>

As the reader can note, the questions investigated in the literature on picture perception get more specific depending on how rich our notion of ‘vision’ is. In this respect, the question about whether we can have, during picture perception, simultaneous conscious or *unconscious* visual representations is, at the moment, the most relevant question in the literature investigating the nature of pictorial experience.<sup>13</sup>

Crucially for the point at stake in this article, the most recent investigation of this question about simultaneity has been offered by Nanay, who has suggested that, while *usual* picture perception (henceforth: UPP) requires that we simultaneously consciously see the depicted object while unconsciously seeing the surface (see also Ferretti 2018a, 2020a, b), *aesthetic* appreciation of pictures, or *aesthetic* picture perception (henceforth: APP) depends on the possibility of consciously seeing, i.e. having a conscious visual experience of, both the depicted object and the surface simultaneously. In this respect, the latter would be visually represented as the vehicle of the pictorial content (Nanay 2011: pp. 461–464; 2015b: p. 192; 2016, 2017: Sect. 2).<sup>14</sup>

If Nanay’s story is true, the possibility of entering a visual state of simultaneous visual consciousness marks the difference between those cases of picture perception in which we have aesthetic appreciation, i.e. APP, and those in which we do not, i.e. UPP: only in APP we simultaneously *consciously* visually represent both the depicted object and the surface.

The reader may be tempted to suppose, at this point, that not only do we already have a complete account of the nature of UPP, as well as of its difference with APP, but also a complete answer to the question about how we simultaneously see, or visually represent, both the surface and the depicted object *consciously and/or unconsciously*. However, things are not so easy.

Recall that a complete answer to the question about simultaneous visual representations in picture perception requires disambiguating the broad sense of ‘seeing’ and of ‘visual representation’, thus specifying whether vision is meant to be conscious or *unconscious*. But, crucially, the same disambiguation is needed when it comes to the definition of ‘conscious’ (and ‘*unconscious*’). Unfortunately, concerning such a disambiguation, the current debate is affected by the same crucial problem that plagued the original dispute between Wollheim and Gombrich: the notions of ‘visual awareness’, ‘visual consciousness’, ‘visual attention’ and ‘visual experience’ used in the literature

<sup>12</sup> Note that, if one does not accept the presence of simultaneous conscious representation, one also needs to specify which component of the picture, either the depicted object or the surface, we visually represent in a conscious manner.

<sup>13</sup> See Hopkins (2012), Briscoe (2018), Nanay (2011, 2015a, 2017), Ferretti (2018a) and Voltolini (2015).

<sup>14</sup> In line with Nanay’s discussion, we use the term ‘consciousness’ as to refer to a subject’s ‘phenomenal experience’: “I treat aesthetic experience as a specific kind of experience and, like in the case of other kinds of experiences, attention plays an important role in determining its phenomenal character” (Nanay 2015b, p. 96). Again, “The general idea is that, to put it simply, aesthetic experiences ‘feel’ different: what it is like to have aesthetic experiences is different from what it is like to have non-aesthetic experiences. The question then is: what is this phenomenal character that is proprietary to aesthetic experiences?” Nanay (2016, p. 20). In the case of vision, phenomenal experience refers to the ‘what it is like’ (Nagel 1974) to see something, sometimes called its phenomenal character. This is what is at stake when investigating the distinction between UPP and APP.

on ‘simultaneity’ are often not clearly defined and, thus, ambiguous.<sup>15</sup> Indeed, the most afflicting issue for the philosophical literature is that the notion of ‘awareness’ is often used as a near synonym of ‘consciousness’, which, in turn, is not properly defined with respect to the notion of ‘attention’, for as we shall see, ‘visual consciousness’ and ‘visual attention’ are sometimes used interchangeably.

To an extent, this use is maintained also in Nanay’s account. For example, when talking about the distinction between attending and not attending, Nanay (2017, footnote 2) explicitly states: “We could also use the conscious/unconscious distinction, as long as we take the unattended stimulus to be unconscious (as we should in the light of empirical evidence (...))”.

From this use, it follows that *attention* and *consciousness* entertain a special relation during visual processing. Although Nanay (2017) acknowledges the possibility that attention is not necessarily conscious, the theoretical shortcut of using these terms interchangeably leads him to overlook the important role of ‘unconscious attentional visual phenomena’ in picture perception. This is, however, very problematic in the light of the specific fact that parallel literatures from vision science and philosophy of perception are more and more sharply defining the differences between ‘visual attention’ and ‘visual consciousness’, the presence of unconscious attention, as well as their interplay in visual perception.<sup>16</sup>

If so, our best theory of picture perception and our best answer to the question about simultaneity have to take into account such a distinction between ‘visual attention’ and ‘visual consciousness’, as well as the role of unconscious attention in visual perception. Otherwise, there is the risk of misconstruing the real role that visual attention plays in constituting the basis of both conscious and *unconscious* visual processing in general and, specifically, during picture perception. Indeed, without a careful distinction between consciousness and attention, a theory of ‘simultaneous visual consciousness in picture perception’ will not offer a rigorous account of how visual consciousness, visual (conscious and unconscious) attention and unconscious visual processing allow the viewer to really distinguish between UPP and APP.

For this reason, in this article we argue that, when it comes to discussing the role of visual attention in picture perception, there is still room to offer a more adequate explanation of the nature of the representational simultaneity at play in both UPP and in APP, with respect to the most promising theories of picture perception we have at the moment. Such an explanation provides a more exhaustive and philosophically satisfying account by explaining the phenomenological peculiarities of pictorial experience in a way that is more faithful to what vision science taught us about the mechanisms of visual attention and their relation to the nature of visual consciousness.

In particular, our account of the nature of picture perception acknowledges, to a full extent, the role of unconscious attentional processes by explaining their relation with conscious visual processing, thus providing a neater demarcation concerning the distinction between UPP and APP, with respect to the role attention plays in both the perception of the surface and of the depicted object.

<sup>15</sup> See Nanay (2011), Ferretti (2018a, b, 2019, 2020a, b) and Hopkins (2012).

<sup>16</sup> For a review see Prinz (2012), Wu (2011), Watzl (2011a, b), Carrasco (2011), Van Boxtel et al. (2010), Ferretti and Marchi (2020).

The most important implication of our account is the following. If attention is not always coupled to consciousness, but it is crucial also for unconscious visual representations, the fact that we are not simultaneously consciously visually representing both the surface and the depicted object does not entail that we are not simultaneously visually attending to both of them. But this means, in turn, that while consciousness is not always simultaneously exercised on both the depicted object and the surface, during both UPP and APP, attention may indeed be so exercised. Therefore, our account suggests a new theory of attention in pictorial perception, which starts from Nanay's proposal, but moves forward: in our view, the exercise of visual attention cannot be the unique crucial factor determining the difference between UPP and APP.

Here is the plan for the article. We first discuss Nanay's view on pictorial attention in more details (Sect. 2) and analyze its problems (Sect. 3). Then, we propose an alternative account (Sect. 4) that does not encounter these problems.

## 2 The *received view* on pictorial attention

As we mentioned, our primary target in this article is the view advocated by Nanay (2017). Nanay offers a theory that tries to explain the way visual consciousness and visual attention operate in picture perception, with respect to the notion of representational simultaneity described above, and with the ambition of clarifying the nature of UPP and APP. This proposal is the most recent and comprehensive attempt to address such problems and, for this reason, we consider it to be the *received view* on pictorial attention. As a starting point, such a view provides a perfect critical target for our discussion.

Nanay's account includes the idea that, in UPP, we visually represent both the depicted object and the surface (Nanay 2011: p. 463; 2017 Sect. 2).<sup>17</sup> This idea is accepted in the literature about UPP (cfr. Sect. 1), as it suggests that, in UPP, we *simultaneously* visually represent, just in a general sense of visually representing, which can be conscious or unconscious, both the surface and the depicted object (Sect. 1). So, this does not entail that we simultaneously *consciously* see both of them. Indeed, this notion is accompanied by a specification (see Sect. 2.2): that, (most of the time) in UPP, we *consciously* visually represent the depicted object, while we *unconsciously* visually represent the surface (2011: p. 463, 2017: Sect. 3; see also Ferretti 2018a).

This construal of UPP can be defended on several grounds, which we cannot review in full details here.<sup>18</sup> The basic idea is that, most of the time, we indeed are conscious of the depicted object, while completely ignoring the surface from the point of view of

---

<sup>17</sup> Nanay sometimes uses the notion of 'awareness' instead of 'attention'. This is because he follows the Wollheimian terminology (1998).

<sup>18</sup> For a review see (Nanay 2011, 2015, 2017; Ferretti 2018a, b, 2020a, b). As anticipated (Sect. 1), the idea that, in UPP, we *must* visually represent, at least *unconsciously*, the surface – while consciously seeing the depicted object – has been recently defended on the basis of the evidence for the perceptual fact that, when we do not (because we cannot) visually represent the surface, even if unconsciously, we enter the illusion that the object is present for real motor interaction, as in the case of pictorial illusions *à la trompe l'oeil* (for a review see Ferretti 2016c, 2018a, 2018b, 2019, 2020a, b).

visual consciousness.<sup>19</sup> One compelling motivation for the claim that, in UPP, we consciously visually represent the depicted object, while we *unconsciously* visually represent the surface, is articulated by Hopkins (2012). According to Hopkins, we *do not* consciously visually perceive the surface, along with the conscious perception of the depicted object, because we *can't*, for the pictorial space we perceive is different from the perceived space occupied by the surface, which is a real space. In this respect, Hopkins (ibid.) notes that if both these perceptions were simultaneously part of our visual conscious experience, picture perception would be a very odd experience, in which several visual-spatial features would clash with each other.<sup>20</sup>

At this point, according to the *received view*, and on the basis of its general commitments to the relation between visual attention and visual consciousness, in UPP, we consciously visually represent and attend to the depicted object, while we unconsciously represent and do not visually attend to the surface. We can summarize these claims about UPP as follows:

- a. We consciously visually represent the depicted object.
- b. We visually attend to the depicted object.
- c. We unconsciously visually represent the surface.
- d. We do not visually attend to the surface.

Importantly, (d) follows from (c), plus the assumption that unconscious visual representations do not require attention. As Nanay (2017) explicitly adds, this should not be problematic, as (p. 166): “Crucially, priming studies show that even unattended objects (like the gorilla) can prime us (that is, it disposes us to be quicker to recognize stimuli that have something to do with gorillas Mack and Rock 1998).” We shall focus on this point and its problems in (Sect. 3, see esp. 3.1). For the moment, let us focus on APP.

From UPP, APP can be construed by modifying the claims (c) and (d) as follows:

- a. We consciously visually represent the depicted object.
- b. We visually attend to the depicted object.
- c1 We consciously visually represent the surface.
- d1 We visually attend to the surface.

Given the above characterization, according to the *received view*, consciousness and attention are always coupled, while it seems that unconscious representations do not need or entail attention (as in UPP). In the next section (see esp. 3.1), we address some major problems with the *received view*, precisely stemming from its characterization concerning the relation between consciousness and attention.

We must acknowledge that Nanay (2016, 2017) endorses a distinction between focal and distributed attention and suggests that, during APP, focal attention concerns the depicted object, while distributed attention pertains to the properties of both the surface and the depicted object. We will discuss this idea below (Sect. 4, see esp.

<sup>19</sup> See Levinson (1998), Lopes (1996: 464), see also the analysis by Nanay (2011: 463, 2015a, 2017) and Ferretti (2018a, b, 2019, 2020a, b).

<sup>20</sup> See also (Nanay 2017) and (Ferretti 2018a, 2019, 2020a, b) for a discussion. Of course, we can sometimes be conscious of the surface while patently ignoring, from the point of view of our consciousness, the depicted object (Nanay 2011; Ferretti 2018a, b). But this is not an interesting case of picture perception.



Sect. 4.2), and suggest an alternative account of how these characteristics of attention (its conscious and unconscious nature, its focal and distributed use) are in play in both UPP and APP in a way that avoids several theoretical problems, we will report, encountered by Nanay’s proposal. Our alternative account has the merit of being more in tune with the empirical evidence from vision science, and also more philosophically consistent.

### 3 Problems for the *received view*

Here we suggest that a commitment to a broad notion of attention that does not capture the important relation between visual attention and conscious or unconscious visual representation is inadequate. Furthermore, as we will show, by focusing on conscious visual attention exclusively, the *received view* offers, at best, only a partial theory of picture perception, which does not fully capture what we know from vision science and from philosophy of perception.

#### 3.1 The problem of assimilation of visual consciousness and visual attention

As we saw in (Sect. 2), in the *received view* the difference between APP and UPP depends on a tight coupling of ‘visual consciousness’ with ‘visual attention’. When one ‘consciously visually represents’ the depicted object, one, ipso facto, also ‘visually attends to’ the object. Conversely, when one ‘unconsciously visually represents’ it, one also ‘lacks visual attention to it’. This commitment determines how we can cash out the main differences between UPP and APP. Nanay explicitly specifies that he uses the notions of ‘attention’ and ‘conscious attention’ interchangeably for a simple reason (2016: p. 474; 2017: Sect. 3): since the debate on the relations between visual consciousness and visual attention is open at the moment, Nanay claims that he does not distinguish between ‘attention’ and ‘conscious attention’ because he does not take any stance on whether attention is necessary and sufficient or, rather, only necessary for consciousness (2011: footnote 4; 2017: Sect. 2). Furthermore, in his writings on pictorial perception and aesthetics, Nanay recognizes the presence of unconscious attention, but does not refer to unconscious attention when analyzing the nature and the differences between UPP and APP, as he states that it is conscious attention that plays a crucial role in his account of (pictorial and non-pictorial) aesthetic experience. So, though the presence of unconscious attention is recognized in his writings, Nanay does not seem to admit any distinctive role for it in picture perception, when explaining the difference between APP and UPP (we will get back to a critical examination of Nanay’s use of these notions with respect to our new account and its technical details in Sect. 4, see esp. Sect. 4.2).<sup>21</sup>

<sup>21</sup> A few additional clarifications about Nanay’s position on unconscious attention and about the fact that he focuses on conscious attention. He writes: “I need to emphasize that what I mean by attention is conscious attention. This is not to deny (or endorse) that attention can be unconscious, but it is conscious distributed attention that I take to be an important feature of the aesthetic domain” (2015b: footnote 6). And: “I need to emphasize that what I mean by attention here and throughout the book is conscious attention. This is very different from the way I have been using the concept of attention in my work in philosophy of perception,



However, as we shall see, this is problematic in the light of empirical results from vision science. Indeed, though the debate concerning the relations between visual attention and visual consciousness is open, there is plenty of evidence that both conscious and unconscious forms of attention play a significant role in one's visual economy, including, as we shall see, the visual representation and appreciation of pictures, something rarely addressed in the literature. This means that also unconscious attention plays a crucial role in visual processing. We need to focus more slowly on this point.

The nature and scope of attentional processes has raised a long-running and intense debate. Different models have been proposed aimed at explaining how attention unfolds, what sort of process it is, and where in the hierarchy of cognitive and perceptual processing it should be located. Recent research has focused on different aspects of attention and has highlighted that various kinds of inputs and tasks can capture a subject's attentional resources (Carrasco 2011: Sect. 1). The kinds of attention that are most widely discussed include, among others, spatial attention, feature-based attention and object-based attention (Chun et al. 2011). The evidence about a multiplicity of attentional phenomena points to a gradual notion of attention, i.e. a process that can be directed to various targets, and located at different stages of cognitive processing.

That said, it is widely accepted in the empirical literature that both conscious and unconscious visual processing require visual attention. The most fundamental reason for accepting such a claim is that human cognition is implemented in a system, the brain, with limited processing resources. As a matter of fact, we know that neural processing has high metabolic costs (Kastner and Ungerleider 2001; Lennie 2003). For this reason, our visual brain needs to be selective toward what it processes, at all stages of processing, at any given time. Researchers have identified the function of attention with the required selection process that filters out information that is deemed not relevant for the system's current task, in order to spare resources (Carrasco 2011; Sect. 2; Watzl 2011a). In other words, attention is often taken to be a mean by which the system optimizes allocation of its limited resources.

This idea has inspired prominent theories of attention such as the early filter view (Broadbent 1958), the late selection view (Deutsch and Deutsch 1963) and, more recently, the biased-competition view (Desimone and Duncan 1995; Duncan 1998, 2006). Thinking of attention as resource optimization in a system with limited

---

Footnote 21 continued

where I allow for unconscious attention (...). I do hold that attention can be conscious or unconscious, but I take conscious distributed attention to be an important feature of the aesthetic domain" (2016: footnote 10). The references to unconscious attention do not suggest that it plays any relevant role in UPP or APP, as, again, the main target is conscious attention: "[...] attending to a property of an object triggers *unconscious* attention to other properties of the same object. But what is at stake in the fourfold distinction I made in the main text is *conscious* attention. While unconscious attention does, as a matter of course, spread to a large number of the properties of the perceived object, we have no reason to believe that conscious attention does" (2016: footnote 13). And even when Nanay recognizes the notion that unconscious perception can rely on unconscious attention, which he takes, however, to be a rather controversial view, he does not consider the importance of unconscious attention for unconscious visual processing during picture perception: "I am assuming in this discussion that 'seeing' here is understood as conscious seeing. I myself am committed to the view that even unconscious perception is systematically influenced by the unconscious allocation of attention, but I am not relying on this (somewhat controversial) view here. I take the claim that conscious perceptual experience depends systematically on the conscious allocation of attention to be not controversial at all (at least not since the inattentive blindness experiments)" (2016: footnote 7).

resources is compatible with the evidence of a multiplicity of attentional mechanisms at different stages of cognitive processing and with different selection-targets. If one considers that the metabolic costs of neural computation remain high throughout the system, it follows that a selection or optimization mechanism could be in place at virtually every stage of processing that occurs in that system (West et al. 2011; Mather and Sutherland 2011; Duncan 2006). This idea is widely accepted in psychological research and suggests that—and this is crucial for the point made in this article—what is not selected by visual attention is not fully processed, or, otherwise the required resources would not be spared (Chun et al. 2011).

If, following the above considerations, visual attention is conceived as resource optimization, we may motivate the following two assumptions:

1. In a system with generally limited processing resources, such as the human (visual) brain, attentional mechanisms are required at virtually all stages of processing.
2. Whatever potential target or property of a target (e.g. spatial location, object features, etc.) is not selected by visual attention, is not fully processed by the visual system.

For the purposes of this article, from these two assumptions we can derive that visual representations require attentional selection mechanisms at every stage.

As we noted, assumption (2) is in line with the tenets of much recent research on attention. However, it is hard to find conclusive empirical evidence that what is not visually attended is not visually processed. At the end of this section, we discuss an experiment on inattentional-blindness that clearly supports this point. For the moment, let us highlight that our second assumption stems from the theoretical considerations we offered above on the resource limitations of the system. Since processing resources are limited, the cognitive system has to be selective toward what is currently processed and what is not.<sup>22</sup> To give a concrete example, the prominent biased-competition model of attention originates from the observation that functional units of the neuronal system respond preferentially to only one stimulus when two or more stimuli fall into their receptive field (Desimone and Duncan 1995). Attentional processes have the role of performing this selection at every stage of the cognitive architecture. This clearly supports assumption (2). Furthermore, if it is indeed hard to find empirical evidence that ‘what is not visually attended is not visually processed’, the same difficulties hold in finding evidence showing that ‘what is visually processed is not attended’.

In this respect, an important additional distinction, to be made here in line with the psychological literature, is the distinction between top-down and bottom-up attention, often referred to as endogenous and exogenous attention respectively (Chun et al. 2011). The notion of top-down (endogenous) attention typically refers to a selection that is operated according to the subject’s goals and task demands. Bottom-up (exogenous) attention is, instead, typically used to refer to attention driven by properties of the

---

<sup>22</sup> This is not to say that attention may be devoted only to one visual item at a time (object, property, spatial location, etc.). Several items could be concurrently selected and allocated processing resources to different degrees. This adds a modulatory role to attention on top of its selective role. However, under the resource limitation assumption, it is still plausible to hold that what is not selected at all is not processed at all.

stimulus.<sup>23</sup> Crucially for our argument, the contention that visual processing requires attention clearly suggests that attention of either kind is necessary for a stimulus to be fully processed, and this holds at any level of visual processing. The distinction between top-down and bottom-up attention, however, does not directly map onto the distinction between conscious and unconscious attention. Indeed, while bottom-up attention is typically thought to work unconsciously and automatically, top-down attention need not always be voluntary and conscious.<sup>24</sup> We believe that attention of both kinds is generally involved in both UPP and APP.

With this distinction in place, one might think that, for what concerns UPP and APP, since the stimulus remains the same, the difference between the two is mainly a difference in top-down attention. In this respect, concerning the relationship between attention and consciousness, a point on which we shall return shortly, it has been claimed that consciousness may arise in the absence of top-down attention. This raises a worry for our proposal: plausibly, being visually conscious of *x* implies that *x* is visually processed. But, if this can happen in the absence of top-down attention, our assumption would be untenable, with respect to the kind of attention relevant for the present discussion.<sup>25</sup>

However, this is not the whole story. Van Boxtel et al. (2010) offer a review of the main line of evidence that there could be (visual) consciousness in the absence of top-down attention. The evidence they report is mainly related to so called ‘*dual-task paradigms*’ in which a subject’s spatial attention is cued to the fixation point within a cognitively demanding task. A target stimulus is presented in the periphery. The subject is typically asked to report about some properties of the peripheral stimulus, e.g. gender, orientation or colour. If the performance of the subject in discriminating the peripheral stimulus is not impaired by the central task, it is then possible to claim that the peripheral stimulus is visually processed and consciously perceived in the absence of attention.

However, what has to be noted here is that these authors are careful in speaking of *near absence* of attention to the peripheral stimulus in the *dual task paradigm*, rather than of complete absence. Indeed, the authors are clearly aware that alternative explanations are available. A very reasonable one is that the central task does not exhaust attentional resources, some of which may still be allocated to the peripheral stimulus. Furthermore, it is the case that different types of attention (e.g. feature-

---

<sup>23</sup> This strict dichotomy has recently been challenged (see, for example, Belopolsky and Thewes, 2012; Thewes, 2019). However, we shall accept it here for the sake of the argument and because it is still widely accepted in psychology.

<sup>24</sup> For example, in their famous experiment on blindsight, Kentridge et al. (1999, experiment c) showed that performance in a visual target detection task would improve for a blindsight subject when presented with an unconscious peripheral cue (i.e. in the blind-field) that would reliably appear at the opposite location from the target, of which the subject was informed. Since the cue would not appear at the same location of the target, the performance increase cannot plausibly be interpreted as an effect of exogenous attention to a spatial location, but rather as an effect of endogenous attention driven by the subject’s knowledge of the cue-location likelihood. This clearly shows that unconscious top-down attentional cues can speed up detection performance, and that top-down attention need not be conscious.

<sup>25</sup> We are grateful to an anonymous referee for asking to technically consider these aspects of attention.

based) may be allocated to the periphery.<sup>26</sup> Thus, empirical evidence of consciousness in the near absence of top-down attention does not suffice to undermine our second assumption and, *ipso facto*, does not represent a problem for our analysis of UPP and APP.

As a further note on this point, van Boxtel et al. (2010) also report evidence of top-down attention in the absence of consciousness (on top of the already discussed evidence for consciousness in the near absence of attention). This implies that, even if the main difference between UPP and APP is a difference in top-down attention, it can still be a difference in unconscious top-down attention.

Now, we agree with Nanay that the matter of the relationship between attention and consciousness is far too wide and controversial to be treated properly in a single article. However, we can still point out that the view that attention is required at all stages of visual processing is in principle compatible with most stances about the relation between attention and consciousness, albeit with different consequences. If attention is necessary and sufficient for consciousness,<sup>27</sup> which is the strongest and most radical among the available positions, what follows from the attentional resource-optimization view outlined above is that every stage of visual processing that requires optimization underlies a conscious state. This view may be considered implausible on several grounds, but there is no inconsistency between the optimization view and the necessity and sufficiency view.

*A fortiori*, there are no inconsistencies between the optimization view and weaker stances on the relation between attention and consciousness. If attention is sufficient but not necessary for consciousness, then the same conclusion as before follows: consciousness embraces each stage of the visual processing hierarchy in which optimization is required. However, on this view, the notion that in systems with unlimited resources consciousness may arise without attention is left open. If attention is necessary but not sufficient for consciousness, from the optimization view it follows that, in the human brain, something other than attention is required to determine at which processing stages consciousness actually arises. In this article, we endorse the latter view, which is supported by studies on blindsight (Kentridge 2004; Kentridge et al. 2008) inattentional blindness (Simons and Chabris 1999) and hemispatial neglect (Driver and Vuilleumier 2001).

On this point, Bressan and Pizzighello (2007) performed a series of five experiment finding, in line with other inattentional blindness studies, that, if engaged in an attention demanding task (e.g. counting the “bounces” of white letters off the borders of a screen), the majority of subjects fail to notice an irrelevant object presented in a display. However, when such an object is present, even if not consciously perceived, performance in the attention-demanding task is reduced. This suggests that attentional

---

<sup>26</sup> Jennings (2015) also expresses similar worries about extant evidence of consciousness in the absence of attention, for gist perception, imagistic consciousness and phenomenal consciousness. However, she suggests a form of consciousness, conscious entrainment, which can happen in the absence of top-down attention. We believe that there are ways to resist the claim that conscious entrainment requires no attention, as it may instead involve an automatization of attentional patterns or attentional habit (Jiang and Sisk, 2019). That said, further discussion of specific aspects of this debate would be beyond the scope of the present paper. The considerations we offered should suffice, for the purposes of the present article, to defuse the worry about the possibility of having any *consciousness without attention*.

<sup>27</sup> As held, for example, by Prinz (2012).

resources are indeed limited, in line with the optimization view, and that some of them are still allocated to the non-consciously perceived object, which despite not surfacing to consciousness, has an impact on the concurrent attention-demanding task. This brings us back to the point concerning the possibility of visual representation without attention. Nanay (2017, p. 166) uses inattention blindness to motivate the claim that vision can occur in the absence of attention, since a subject may be primed by a stimulus that is not consciously perceived when attention is engaged in a different task. However, the experiment we have just presented suggests a different interpretation. Indeed, the stimulus appears to be visually processed by the subject, as it exerts a priming effect. But this is so because some attentional processing resources are detracted from the central task, which is clearly shown by the performance interference. Finally, this seems not to be enough to make the stimulus conscious, this being in line with the necessity, but not sufficiency, of attention for consciousness.<sup>28</sup>

Research on attention, as we have seen, suggests that both conscious and unconscious visual processing compete for the same attentional resources. If this is correct, then there are serious reasons to doubt that in UPP we do not visually attend to the surface (Sect. 2). At the very least, thus, the proponent of the *received view* has to offer additional support for the claim that there is no (unconscious top-down) attention exercised on the surface in UPP. Otherwise (d) cannot be convincingly maintained anymore. But if one abandons (d), the distinction between UPP and APP becomes weaker, as in both of them we are committed to (d1): we attend to the surface. Thus, the distinction rests only on the difference between (c) and (c1). As we shall see, the difference would be that, in UPP, we *unconsciously* visually represent the surface, while in APP we *consciously* visually represent the surface. Having that said, as we will propose with important specifications, it is possible to show that in both of them we attend to the surface, but stating that we unconsciously attend to the surface in UPP, while consciously attending to the surface in APP. We will come back to this point in (Sect. 4). Before doing so, we need to address another problem for the *Received View*.

### 3.2 The problem of an odd visual experience in APP

As we saw, the *received view* endorses a special equation between consciousness and attention. The account also seems to embrace Hopkins' view and the related worries about conscious simultaneity (cfr. Sect. 2) and, thus, it wants to avoid the problems related to the claim about simultaneity of visual consciousness of both the surface and the depicted object (Nanay 2017: Sect. 2).

An important point is the following: Hopkins (2012) only talks about the notion of 'visual experience'. And it is reasonable to suppose that what Hopkins has in mind, like most philosophers, is 'conscious visual experience', i.e. visual consciousness—and, arguably, not attention, which is never mentioned. In this respect, Nanay observes

<sup>28</sup> The view that attention is necessary but not sufficient for consciousness has the advantage of allowing for a constraint on the scope of consciousness in the processing hierarchy, which seems plausible when taking into consideration both neurophysiological and behavioral data (Koch, 2004; Boyer et al. 2005), as well as evidence related on introspection (e.g. disaggregated object features processed at early levels of visual processing do not seem to figure among the introspectively available contents of conscious experience). See also (Jackendoff 1987; Prinz 2012).

that Hopkins' worry holds only if by 'seeing' we mean 'having visual experience', this being equated with 'having visual consciousness'. But in the received view, visual consciousness is equated to visual attention, to 'visually attending to' (Ibid.). Thus, if consciousness and attention are equated, Hopkins's worry arises with 'simultaneous visual attention', this being equated to 'simultaneous visual consciousness'.

For this reason, the *received view* holds that, during UPP, we consciously visually experience and attend to the depicted object, while unconsciously seeing the surface, this implying exercising no attention on it (cfr. Sect. 2). Indeed, according to the *received view*, there seems to be no simultaneous consciousness, or conscious attention, during UPP:

if we are simultaneously attending to both the depicted scene and the picture's surface, then there seems to be something contradictory or disjoint about our simultaneous experience of both of these. But, crucially, this objection does not apply if pictorial twofoldness is understood not as simultaneous attention, but as simultaneous (conscious or unconscious) representations (Nanay 2015a, 192; see also 2017: Sect. 2).

This move apparently saves Nanay's account from Hopkins' argument. Why apparently? Because in the same pages in which Nanay considers the objection to UPP and, indeed, also tries to take into account it when analyzing UPP, he does not consider, however, this problem when analyzing APP. Indeed, he points out that:

We have seen that one way of making the proposal about simultaneous seeing work when it comes to understanding picture perception (not appreciation) is to bring in the concept of attention and to argue that while we do simultaneously see both the surface and the depicted scene, we do not simultaneously attend to both—we are only attending to the latter. But those special cases in which we are aesthetically appreciating pictures are different. Then, in addition to simultaneously seeing both the surface and the depicted scene, we also attend to the surface and the depicted scene simultaneously. Each time we see something in a picture, we see both the surface and the depicted scene. We can attend to either—although we normally attend to the latter only. But we can direct our attention to the picture surface as well as to the relation between the two. And this is what happens when we appreciate pictures aesthetically. The aesthetic appreciation of pictures is a form of picture perception where our attention is exercised in a special manner. (Nanay 2017: p. 7).

The conclusion is that, following the *received view*, UPP is different from APP in that, "in order to appreciate a picture aesthetically, one needs to exercise twofold attention: attending to both the picture surface and the depicted object" (2017: p. 8), as we reported in (Sect. 2).

The shrewd reader may note, however, that such a construction of APP constitutes another problem for the *received view*, precisely in the light of Hopkins' point (cfr. Sect. 2): 'simultaneous visual attention', this being equated, by the *received view*, to 'simultaneous visual consciousness' to both the depicted object and the surface, during APP, would lead to an odd visual experience. This means that, if we enter such a

simultaneity during APP, then APP should lead to an odd visual experience—we'll come back to this point more carefully in (Sect. 4).

Since, as noted, when Nanay talks about 'simultaneous attention', he means 'simultaneous conscious attention', as 'attention' and 'consciousness' are equated, the *received view* cannot hold, at the same time, (i) that Hopkins is right in denying simultaneity of conscious visual experience, and, thus of visual attention to, both the surface and the depicted object in UPP, and (ii) that we simultaneously attend to both of them in APP, this implying, *contra* Hopkins, that we simultaneously consciously visually experience (and attend to) both of them in this second case. If so, *the received view* falls prey of the problem raised by Hopkins (2012) about the 'oddity' of visual experience, arising from the commitment to 'simultaneous visual attention', which is equated to 'simultaneous visual consciousness', when it comes to a definition of APP.

In the next section, we offer an alternative account that is capable of facing the two problems afflicting the *received view* outlined in this section.

## 4 Beyond the *received view*

We saw that Nanay's account faces two main problems, which are related.

1. Both conscious and unconscious visual processing require attention. This is a problem for the definition of UPP, as well as for accurately describing the difference between APP and UPP (Sect. 3.1).
2. The problem of oddity in 'simultaneous visual consciousness' raised by Hopkins (Sect. 3.2), concerning the explanation of APP.

Here we offer an account that does not face these problems. Let us start with (1).

### 4.1 Attention, consciousness and UPP

In (Sect. 3.1), we saw that attention is always involved in both conscious and unconscious visual processing and that this is a problem for the *received view*. Indeed, there are several reasons to seriously doubt that, in UPP, (d) is the case, i.e. that we do not visually attend to the surface (Sect. 2). At this point, the definition of UPP can be renewed.

Indeed, the reader should note that we accept the idea about simultaneity proposed by the *received view*, and the literature (Sect. 2), that in UPP we consciously see the depicted object while we do not need to consciously see the surface. However, drawing on evidence that both conscious and unconscious visual processing require attention, we propose a complementary view concerning the role played by attention in picture perception, which, differently from the *received view*, posits a role for visual attention even in UPP. The view we obtain is the following.

In UPP, we always need to have visual attention attuned to both the surface and the depicted object, even if only the latter is consciously visually perceived and attended, while the former is attended, but unconsciously visually perceived (i.e. it is *unconsciously attended*). We can go more slowly on this point. First:



- (i) Both conscious and unconscious visual representations require visual attention—from (Sect. 3.1).

Second, in UPP:

- (ii) We consciously visually represent the depicted object – in line with the *received view*, from (Sect. 2).

And:

- (iii) We visually represent, *unconsciously*, the surface—(c), from (Sect. 2).

From the conjunction of these three points, we have that, *contra* the *received view*, in UPP:

- (iv) We visually attend to both the depicted object and the surface. However, only the former is consciously visually represented. The latter is unconsciously visually represented. In other words, we consciously see and attend to the depicted object, while we unconsciously see and attend to the surface.

We saw that (ii) and (iii) are accepted by the *received view*, as well as in the literature on UPP, and are compatible with the results from vision science. So, as the reader can see, (i) sets the first difference between our view and Nanay's. If (i) is true, as suggested by research on attention (Sect. 3.1), then it implies the falsity of (d) and, thus, of the *received view*'s construal of UPP.<sup>29</sup>

As we saw, (Sect. 3.1), the rejection of (d) leads to the conclusion that the distinction between UPP and APP is weaker. Indeed, in both of them, we are committed to (d1): we attend to the surface. Thus, the difference between APP and UPP rests only on the difference between (c)—i.e. we unconsciously visually represent the surface—and (c1)—we consciously visually represent the surface. In other words, the difference is the following: in UPP, we unconsciously visually represent the surface, while in APP we consciously visually represent the surface.

At this point, one may argue that, while our move includes the results from vision science in our philosophical theory of picture perception, as to avoid (d) in the light of (i), and this leads to a renewal of the definition of UPP, we still need to describe our commitment to the idea that, in APP, we consciously visually represent both the depicted object and the surface. Indeed, this is a problem in the light of Hopkins' objection, which constitutes a serious shortcoming of the *received view*, and which a new theory should be able to overcome. In the next section, we suggest how our account can avoid this problem.

<sup>29</sup> We need to acknowledge that, for its philosophical purposes, the *received view* exemplified by Nanay's book (2016) explicitly focuses on conscious attention, without addressing the role of unconscious attention in picture perception. His article follows this methodological stance (2017) as well. That said, and though we are not against this methodological decision, it remains true that not considering the role and the nature of unconscious attention in picture perception rules out several important differences between the APP and UPP. Thus, it seems we cannot avoid taking unconscious attention into consideration. Starting from Nanay's account, and being aware of these considerations, our move wants to fill the gap concerning the presence of unconscious attention and its peculiar role in picture perception.

## 4.2 Why APP is not odd

We suggested that vision always requires attention. Thus, *ipso facto*, and *contra* the received view, in UPP, as well as in APP, the visual representation of the surface always requires attention. However, as anticipated in (Sect. 3.2), another problem for the *received view* is to account for APP without falling victim of Hopkins' objection. Since our renewal of the notion of UPP does not free us from this problem, as, from the *received view* we inherit the idea that, in APP, we consciously visually represent both the depicted object and the surface, here we need to explain how our account can overcome this problem of accurately defining the nature of APP without falling prey to Hopkins' objection.

We have proposed to distinguish between visual attention and visual consciousness. We have also assessed the way in which attention is responsible for both conscious and unconscious visual representations. But this is not the whole story, if we aim to meticulously follow what vision science has taught us about the mechanisms of attention. We can, indeed, further distinguish between two kinds of visual attention: *focal* and *distributed* attention. Both these forms of visual attention can be conscious. We already introduced the widespread assumption that attention optimizes the allocation of limited processing resources (Sect. 3.1). Now, these resources can be allocated to the same visual object, as well as to several visual objects in the visual field. In case of all resources devoted to only one object, attention is *focal*. In case of resources devoted to several different objects, attention is *distributed* (De Brigard and Prinz 2009; Cohen and Dennett 2011; Eriksen and Hoffman 1972; see also Ferretti and Marchi 2020). However, and crucially for the point at stake in this paper, visual attention can be focused on one property or distributed to several properties of the same or different objects - as also Nanay (2015b) recognizes.

If so, it follows that focal and distributed attention can be both in play during the same visual experience, and flow seamlessly into one another. To understand how, think about fireworks. Before the show starts, the spectator is looking and visually attending to an extended and roughly delimited spatial location. In this case, attention is *distributed* to the whole area. Once the show starts, however, single shots will initially be fired and the visual attention devoted to each of them will be *focal*. Towards the end of the show, multiple shots will be fired at the same time and visual attention will, again, be *distributed* to several of them at the same time. In this example, visual attention optimizes allocation of resources in a different manner: firstly, to an entire extended spatial location; secondly, to individual items at that location; thirdly, to multiple items at that location. Given the limited processing resources constraint (Sect. 3.1), change in optimization will correspond to changes in visual resolution and detail in the three conditions.

Accordingly, we know that, in distributed attention, targets are detected or experienced in less rich details or resolution (Cohen and Dennett 2011), whereas in certain cases of focal attention, very salient targets that are outside of the focus are not even detected (Simons and Chabris 1999). Furthermore, a crucial thing to note is that, in distributed attention, it is not the case that limited resources must always be distributed equally among all the targets of attention. Even if attention is distributed to more than

one object, or property, some of these objects, or properties, may still be experienced in greater resolution. How does this further analysis of visual attention apply to our definition of APP in a way that leads us to overcome Hopkins' problem? Well, we just need to add more details to the notion of visual attention we are coupling to visual consciousness.

We saw that simultaneous visual consciousness of both the surface and the depicted object would lead to an odd visual experience, according to Hopkins' worry. Should we give up with the idea that, in APP, we consciously attend to both the surface and the depicted object (as one might argue that this could give rise to an odd visual experience), or is there a way for us to maintain it without problems? We think that the distinction between distributed and focal attention can help us to maintain such a definition of APP, while also permitting to avoid Hopkins' worry. And this constitutes a big advantage of our account.

As we saw in (Sect. 3.2), Hopkins talks about simultaneous visual experience. So, neither he mentions any form of visual attention, nor he mentions any degree of visual consciousness. Here is, then, a more accurate description of APP. In the case of APP, it is possible to *simultaneously consciously attend* to both the surface and to the depicted object. But, *crucially*, this is possible only if the visual attention we are talking about is *distributed conscious attention*. Of course, in this case, the resolution of conscious visual experience is slightly degraded. This is due to the limited capacities of attentional resources. However, this is what opens to a crucial perceptual possibility in the case of APP (Nanay 2017: p. 7, see also Sects. 3 and 4): consciously attending not only to the depicted object and the surface, but also to the relation between the two (Sect. 3.2). This is consistent with the idea that, indeed, in the case of APP, there is a special attunement of our visual system to the two components of the picture, the depicted object and the surface, which is not normally achieved in UPP.

This special attunement is given by the fact that distributed conscious attention relates to both the surface and the depicted object. Why such an attunement does not lead to an odd visual experience? This is precisely because, in APP, conscious attention is *distributed*, not *focal*.<sup>30</sup> And this is perfectly in line with the idea by Hopkins that we cannot really have a full conscious visual experience of both the surface and the depicted object, with respect to their different spatial properties. And this is because we cannot have conscious *focal* attention attuned to both. Therefore, simultaneous visual consciousness of both the depicted object and the surface would lead to an odd visual experience of overlapping visual-spatial features only if we are talking about '*simultaneous conscious focal attention*'. But this cannot happen because what we can rely on is, at best, *distributed* conscious attention.

It is worth noting that *conscious focal attention* can alternate between the surface and the depicted object. And, in this case, there is no lack of resolution, as conscious focal attention can be exercised on one target at a given time, and not simultaneously. Again, the possibility of relying on conscious focal attention simultaneously attuned to both the surface and the depicted object, would lead us to have an odd visual experience, as suggested by Hopkins. However, we never enter such an odd visual experience—we

<sup>30</sup> This point demarcates our view from Nanay's one (2016), who also talks about distributed attention, but acknowledges that attention can be both focused and distributed at the same time, which we exclude. Further discussion below.

cannot even try to do so—because this is not possible given to resource limitations at the basis of our visual attentional processes.<sup>31</sup> This is also in line with the fact that, in APP, we can visually experience the depicted object that is encoded within the surface, as well as perceptually appreciate that its visual attributes are related to the properties of the design as design. In this case, thus, not only do we see the surface, but we also enter design-seeing: we appreciate the properties of the surface that are responsible for the way in which the depicted object is (re-)presented. In this respect, mere surface-seeing permits the viewer to enter UPP and, thus, not to fool one about the object's real presence (for a review of this point, see Ferretti 2016c, 2018a, b, 2020a). This is in line with the fact that mere surface seeing can be unconscious (Nanay 2010b, 2011, 2017; Lopes 2005; Voltolini 2013; Hopkins 2010; Ferretti 2016c, 2018a, 2018b, 2020a, b). However, design-seeing permits to enter APP and to appreciate the relation between the design properties from which the pictorial properties of the depiction emerge, or are visually encoded (Nanay 2010b, 2011, 2017; Lopes 2005; Voltolini 2013; Hopkins 2010; Ferretti 2018a).

So, our theory of visual attention in pictorial perception is safe, as, thanks to the possibility of invoking the role played by *conscious distributed attention*, we can explain why it is that we do not enter a disjointed and odd visual experience in APP. Furthermore, such a form of visual attention also allows us to appreciate, in APP, how the visual characteristics of the depicted object emerge from those pertaining to the surface.

Summing up, our explanation of the difference between conscious *focal* and *distributed* attention allows us to maintain the definition of APP according to which, differently from the UPP, we consciously attend to both the surface and the depicted object, while, at the same time, avoiding Hopkins' worry: in APP we can only have simultaneous *distributed* conscious attention exercised on the surface and the depicted object. If so, our account also explains why Hopkins' worry is confirmed by vision science.

But there is another important point. We introduced the venerable dispute between Wollheim (1987, 1980, 1998), who defended the idea of *simultaneity*, and Gombrich (1960), who argued for the *alternation* of these visual states, as well as the way such a debate has been pushed even further in contemporary reflections on picture perception.<sup>32</sup> In this respect, our definition of APP paves the road for a renewal of the debate between alternation and simultaneity. We now see that alternation and simultaneity are not just about 'seeing' in general terms (Sect. 2), as well as that they are not only about 'consciousness' in general terms (Ibid.). They can be, more specifically, about visual attention. In particular, about how visually attention can be *distributed* or *focal*. This deepens our story about pictorial perception, explaining how

<sup>31</sup> Recall that it has been recently suggested that, in line with the view that, in UPP, we must simultaneously visually represent both the surface and the depicted object, but we do not consciously see both of them: we can either consciously see the depicted object while unconsciously seeing the surface, or consciously see the surface, while unconsciously seeing the depicted object (for recent reviews see, Ferretti 2018a, b, 2020a). Our account can extend this view by suggesting that, in APP, we can exercise not only consciousness in general, but *conscious focal attention* only on one of these two components of the picture at a given time.

<sup>32</sup> Cfr. footnote 8.

its nature is much more complex, and describing simultaneity in a more accurate and technical manner.

Now, as anticipated (Sect. 3.1), we must acknowledge that the distinction between *focal* and *distributed* attention is already present in the *received view*. Nanay (2016) argues that, in pictorial experiences, attention is focused on one object and distributed to several properties of that object,<sup>33</sup> for *focal* attention concerns the depicted object, while attention pertains to the properties of the surface. However, our account is beneficial because it points out something about pictorial attention which is not mentioned in Nanay's account, and which is crucial for having a coherent account of APP and UPP. Let us go more slowly on this.

First, Nanay does not aim to offer a comprehensive account that explains the relation between conscious and unconscious visual attention, as well as of the relation between *focal* and *distributed* attention, when it comes to defining the difference between UPP and APP. This is, however, what we aim to do here. Indeed, Nanay does not focus on the role of unconscious attention in pictorial perception, and despite his mention of both focal and distributed attention, his account heavily relies on the notion of conscious attention, and in particular of conscious distributed attention, not focal (cfr. the quotes reported in fn 21). This makes his account very different from the theory proposed in the present paper.

Second, Nanay does not explicitly apply the distinction between focal and distributed attention as a solution to Hopkins' worry, as we do here—thus, his account falls prey of this big problem.

Furthermore, it is not the case that, as suggested by Nanay, in APP we have focal attention exercised on the depicted object and distributed attention exercised on the surface. Indeed, once we have distributed attention to one visual scene (object plus surface) we cannot have, at the same time, focal attention on a special portion of it. Conversely, when we have focal attention, we cannot have distributed attention exercised on other parts of the scene. Curiously, this latter point seems to be supported by findings reported by Nanay himself (2015b: p. 21) about different looking patterns on the same images between advanced art-school students and artistically untrained subjects: experts have much more distributed attention when looking at pictures when compared to naïve observers (Vogt and Magnussen 2007). The idea that such results show a subject's attentional deployment is controversial, as Nanay himself notices. Attention can be exerted covertly, i.e. without moving one's eyes, and these results are, at most, relevant to a subject's overt attentional pattern. Furthermore, it is not clear whether these patterns really reflect aesthetic vs. non-aesthetic experience of the picture. But assuming, for the sake of the argument, that a subject's attentional pattern, related to aesthetic experience, is really reflected in these measurements, the significantly different looking patterns, where the naïve observer's pattern is limited

---

<sup>33</sup> Nanay basically defends the claims about APP exposed in his (2015b) and in his (2017a) even in his new book (2017b, see Sect. 3, esp. Sect. 3.4). Note that someone has recently criticized Nanay's general view about aesthetic attention (Fazekas 2016). However, this critic does not aim to take into account Nanay's theory of depiction. Differently, here we just aim to propose our view about the importance of unconscious attention in picture perception and, *ipso facto*, we suggest how we should conceive the difference between normal and aesthetic pictorial experience. But this is not in contrast with Nanay's overall view of aesthetic experience.

to a small portion of the image and the expert's pattern covers most of the image, strongly supports the view that attention is *either* focal or distributed.

So, here we propose that, in APP, one only has distributed attention exercised on both the surface and the depicted object. This point differentiates our view from Nanay's one. While the amount of resources that are '*distributed*' to the surface and the depicted object may be different, in APP one target, namely the depicted object, can still be visually experienced in greater resolution than the surface. However, neither will be experienced with the same resolution as when focal attention is deployed to one or the other. This is in line with Nanay's (2010a) earlier view about visual attention modulating the degree of determination of the properties we visually perceive.

Summing up, even if the distinction between distributed and focal attention and its role in aesthetic experience is already partly addressed by the *received view*, some aspects of the relation between focal and distributed attention were not clear at all, and such a distinction was not employed as a possible solution to the problem of the odd visual-experience raised by Hopkins. In this respect, our view takes into account these two points and, in doing so, it offers a new theory of pictorial attention, which represents an improved version of the received view, and which is now completely immune to these problems.

We want to conclude by discussing how our proposal that in UPP one is always deploying unconscious attention to the surface may be further supported. First, it has been suggested that visual attention has the role of guiding action (Wu 2011; Watzl 2011a, b). Second, as we have seen, in UPP we need to visually represent, at least *unconsciously*, the surface as an object for motor interaction, because, when we cannot visually track the presence of the surface, we are fooled, at the conscious level, that the depicted object is a present object we can interact with, as in the case of *trompe l'oeil* pictorial illusions (we do not discuss this case here, for a philosophical review see Ferretti 2016b, 2017b, 2018a, b, 2019, 2020a, b; Briscoe 2016 2018; see also Vishwanath 2011, 2014, Vishwanath and Hibbard 2010, 2013). This is because the unconscious perception of the surface can modulate our conscious perception of the depicted object (Nanay 2017; Sect. 2; Ferretti 2016b, 2017b, 2018a, b, 2020a, b). If we couple this second notion with the first point that visual attention has the role to guide action and with the claim that all visual processing needs attention, we cash out the idea that unconscious visual attention is crucial for us to visually represent, at least *unconsciously*, the surface as an object for motor interaction. Without unconscious attention, we would fall into the pictorial illusion of presence of the depicted object, as in the case of *trompe l'oeils*.

Second, our view permits to avoid what is called the '*refrigerator light illusion*' also in the case of picture perception. According to the '*refrigerator light illusion*', "we are misled, because we become conscious of something as soon as we focus on it, just like someone might naively think that the refrigerator light is always on because it is on as soon as he looks" (Watzl 2011b: p. 723). Accordingly, in picture perception, we think that we visually consciously attend to the surface only when we are visually conscious of it. But, as we suggested, this is not the case: we always project our unconscious visual attention onto the surface. Thus while we do not always need to simultaneously visually experience the depicted object and the surface at the same time, during picture perception, we can divide our visual attention between them

(see also Briscoe 2018: 76; Voltolini 2013, 2015 for a similar point) All this seems to suggest that our claim that even UPP needs visual attention exercised on both the surface and the depicted object is a crucial claim about the nature of UPP. A claim not considered by Nanay, but that can still be in tune with his account.

## 5 Conclusion

We discussed the *received view* on picture perception and its problems. The two basic problems encountered by the *received view* are the following: such a view does not adequately consider the role of unconscious attention in picture perception, and it does not explain how we can reach APP without entering an odd visual experience, as described by Hopkins.

We proposed an alternative account capable of overcoming these two problems. For this reason, our account is beneficial for the current debate on picture perception, as it offers a new perspective on pictorial attention. Our theory, indeed, is the first which, at once, (a) explains the complete role of visual attention (and makes justice to the role of unconscious attention) in picture perception; (b) does that by considering visual attention in all its complexity (conscious and unconscious, focal and distributed, top-down and bottom-up) and by considering its relation with visual consciousness; (c) provides a coherent account of UPP, of APP, as well as of their respective visual differences in the light of such a complexity; and (d) explains how consciousness and attention work during APP, without, however, leading the viewer to an odd visual experience.<sup>34</sup>

## References

- Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology*, 2, 1–8. <https://doi.org/10.3389/fpsyg.2011.00246>.
- Boyer, J. L., Harrison, S., & Ro, T. (2005). Unconscious processing of orientation and color without primary visual cortex. *Proceedings of the National Academy of Sciences*, 102(46), 16875–16879. <https://doi.org/10.1073/pnas.0505332102>.
- Bressan, P., & Pizzighello, S. (2008). The attentional cost of inattention blindness. *Cognition*, 106(1), 370–383. <https://doi.org/10.1016/j.cognition.2007.03.001>.
- Briscoe, R. (2016). Depiction, pictorial experience, and vision science. In C. Hill & B. McLaughlin (Eds.), *Philosophical topics, special issue on appearance properties*, Vol. 44(2). Fayetteville: University of Arkansas Press.

<sup>34</sup> Both authors contributed equally to the arguments and the writing of this article. The order of authors is purely alphabetical. Gabriele Ferretti was supported by a NOMIS Fellowship, awarded by the Eikones — Center for the Theory and History of the Image at the University of Basel, Switzerland. Francesco Marchi was supported by the Mercator Research Center Ruhr (MERCUR) project Pr-2016–0016 and by the Center for Mind and Cognition of the Ruhr Universität Bochum. We would like to warmly thank two anonymous reviewers for addressing important comments, both on the empirical and on the philosophical side, which led us to significantly improve the first version of this article. We also thank Albert Newen, as the idea for this paper was born as a beautiful result of a quick unexpected conversation in front of a blackboard between the authors, during a visiting fellowship of Gabriele Ferretti, hosted by Newen, at the Ruhr Universität Bochum. Finally, we are grateful to Bence Nanay and Albert Newen, for discussing with us the ideas offered in this paper.



- Briscoe, R. (2018). Gombrich and the duck-rabbit. In M. Beaney (Ed.), *Aspect perception after Wittgenstein: Seeing-as and novelty*. Abingdon: Routledge.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.
- Cattaneo, Z., & Vecchi, T. (2011). *Blind vision: The neuroscience of visual impairment*. Cambridge: MIT Press.
- Cavedon-Taylor, D. (2011). The space of seeing. *British Journal of Aesthetics*, 51(3), 271–278.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <https://doi.org/10.1016/j.visres.2011.04.012>.
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62(1), 73–101. <https://doi.org/10.1146/annurev.psych.093008.100427>.
- Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, 15(8), 358–364. <https://doi.org/10.1016/j.tics.2011.06.008>.
- De Brigard, F., & Prinz, J. (2010). Attention and consciousness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1), 51–59. <https://doi.org/10.1002/wcs.27>.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 70, 80–90.
- Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition*, 79(1–2), 39–88.
- Duncan, J. (1998). Converging levels of analysis in the cognitive neuroscience of visual attention. *Philosophical Transactions of the Royal Society of London, Series B-Biological Sciences*, 353, 1307–1317.
- Duncan, J. (2006). EPSmid-career award 2004: Brain mechanisms of attention. *The Quarterly Journal of Experimental Psychology*, 59(1), 2–27.
- Eriksen, C. W. W., & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics*, 12, 201–204.
- Fazekas, P. (2016). Attention and aesthetic experience. *Journal of Consciousness Studies*, 23(9–10), 66–87.
- Ferretti, G. (2016a). Pictures, action properties and motor related effects. *Synthese Special Issue Neuroscience and its Philosophy*, 193(12), 3787–3817. <https://doi.org/10.1007/s11229-016-1097-x>.
- Ferretti, G. (2016b). Visual feeling of presence. *Pacific Philosophical Quarterly*. <https://doi.org/10.1111/papq.12170>.
- Ferretti, G. (2017a). Pictures, emotions, and the dorsal/ventral account of picture perception. *Review of Philosophy and Psychology*, 8, 595–616. <https://doi.org/10.1007/s13164-017-0330-y>.
- Ferretti, G. (2017b). Are pictures peculiar objects of perception? *Journal of the American Philosophical Association*, 3(3), 372–393. <https://doi.org/10.1017/apa.2017.28>.
- Ferretti, G. (2017c). Two visual systems in molyneux subjects. *Phenomenology and the Cognitive Sciences*, 17(4), 643–679.
- Ferretti, G. (2018a). The neural dynamics of seeing. *Erkenntnis*, 84(6), 1285–1324. <https://doi.org/10.1007/s10670-018-0060-2>.
- Ferretti, G. (2018b). The nature of pictorial representations. *Phenomenology and Mind*, 4, 136–144. [https://doi.org/10.13128/Phe\\_Mi-23631](https://doi.org/10.13128/Phe_Mi-23631).
- Ferretti, G. (2019). Perceiving surfaces (and what they depict). In B. Glenney & J. F. Silva (Eds.), *The senses and the history of philosophy*. Abingdon: Routledge.
- Ferretti, G. (2020a). Why *Trompe l'oeils* deceive our visual experience. *The Journal of Aesthetics and Art Criticism*, 78–1, 33–42.
- Ferretti, G. (2020b). Do *Trompe l'oeils* look right when viewed from the wrong place? *The Journal of Aesthetics and Art Criticism*, 78–3, 319–330.
- Ferretti, G., & Glenney, B. (2020). Molyneux's question and the history of philosophy. Routledge.
- Ferretti, G., & Marchi, F. (2020). Seeing entities without seeing N-entities. *Journal of Consciousness Studies*, 27(1–2), 57–70.
- Gombrich, E. (1960). *Art and illusion*. New York: Pantheon.
- Hopkins, R. (2003). Pictures, phenomenology and cognitive science. *Monist*, 86, 653–675.
- Hopkins, R. (2010). Inflected pictorial experience: Its treatment and significance. In C. Abell & K. Bantilaki (Eds.), *Philosophical perspectives on depiction*. Oxford: Oxford University Press.
- Hopkins, R. (2012). Seeing-in and seeming to see. *Analysis*, 72, 650–659.
- Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge: MIT Press.

- Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia*, *39*(12), 1263–1276.
- Kentridge, R. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia*, *42*(6), 831–835.
- Kentridge, R. W., Nijboer, T. C. W., & Heywood, C. A. (2008). Attended but unseen: Visual attention is not sufficient for visual awareness. *Neuropsychologia*, *46*(3), 864–869.
- Koch, C. (2004). *The quest for consciousness*. Boston: Roberts Publishers.
- Kulvicki, J. (2006). *On images: Their structure and content*. Oxford: Oxford University Press.
- Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497. [https://doi.org/10.1016/S0960-9822\(03\)00135-0](https://doi.org/10.1016/S0960-9822(03)00135-0).
- Levinson, J. (1998). Wollheim on pictorial representation. *Journal of Aesthetics and Art Criticism*, *56*, 227–233.
- Lopes, D. M. (1996). *Understanding pictures*. Oxford: Oxford University Press.
- Lopes, D. M. (2005). *Sight and sensibility. Evaluating pictures*. Oxford: Oxford University Press.
- Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133.
- Matthen, M. (2005). *Seeing, doing and knowing: a philosophical theory of sense perception*. Oxford: Oxford University Press.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450.
- Nanay, B. (2005). Is twofoldness necessary for representational seeing? *British Journal of Aesthetics*, *45*(3), 248–257.
- Nanay, B. (2010a). Attention and perceptual content. *Analysis*, *70*(2), 263–270.
- Nanay, B. (2010b). Inflected and uninflected experience of pictures. In C. Abell & K. Bantilaki (Eds.), *Philosophical perspectives on depiction*. Oxford: Oxford University Press.
- Nanay, B. (2011). Perceiving pictures. *Phenomenology and the Cognitive Sciences*, *10*, 461–480.
- Nanay, B. (2015a). Trompe l’oeil and the dorsal/ventral account of picture perception. *Review of Philosophy and Psychology*, *6*, 181–197.
- Nanay, B. (2015b). Aesthetic attention. *Journal of Consciousness Studies*, *22*(5–6), 96–118.
- Nanay, B. (2016). *Aesthetics as philosophy of perception*. Oxford: Oxford University Press.
- Nanay, B. (2017). Threefoldness’. *Philosophical Studies*. <https://doi.org/10.1007/s11098-017-0860-2>.
- Newall, M. (2011). *What is a picture?*. Basingstoke: Palgrave Macmillan.
- Prinz, J. J. (2012). *The conscious brain: How attention engenders experience*. Oxford: Oxford University Press.
- Ro, T., & Rafal, R. D. (1996). Perception of geometric illusions in hemispatial neglect. *Neuropsychologia*, *34*(10), 973–978.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattention blindness for dynamic events. *Perception*, *28*(9), 1059–1074.
- van Boxtel, J. J., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: On sufficiency and necessity. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2010.00217>.
- Vishwanath, D. (2011). Information in surface and depth perception: Reconciling pictures and reality. In L. Albertazzi, G. J. van Tonder, & D. Vishwanath (Eds.), *Perception beyond inference. The information content of visual processes* (pp. 201–240). Cambridge: MIT Press.
- Vishwanath, D. (2014). Toward a new theory of stereopsis. *Psychological Review*, *121*(2), 151–178.
- Vishwanath, D., & Hibbard, P. (2010). Quality in depth perception: The plastic effect. *Journal of Vision*. <https://doi.org/10.1167/10.7.42>.
- Vishwanath, D., & Hibbard, P. (2013). Seeing in 3D with just one eye: Stereopsis in the absence of binocular disparities. *Psychological Science*, *24*, 1673–1685. <https://doi.org/10.1177/0956797613477867>.
- Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye-movement patterns and visual memory in artists and laymen. *Perception*, *36*(1), 91–100.
- Voltoini, A. (2013). Why, as responsible for figurativity, seeing-in can only be inflected seeing-in. *Phenomenology and the Cognitive Sciences*, *14*(3), 651–667. <https://doi.org/10.1007/s11097-013-9335-x>.
- Voltoini, A. (2015). *A syncretistic theory of depiction*. Basingstoke: Palgrave Macmillan.
- Watzl, S. (2011a). The nature of attention. *Philosophy Compass*, *6*(11), 842–853. <https://doi.org/10.1111/j.1747-9991.2011.00433.x>.
- Watzl, S. (2011b). The Philosophical Significance of Attention. *Philosophy Compass*, *6*(10), 722–733.

- West, G. L., Anderson, A. A. K., Ferber, S., & Pratt, J. (2011). Electrophysiological evidence for biased competition in V1 for fear expressions. *Journal of Cognitive Neuroscience*, 23(11), 3410–3418.
- Wollheim, R. (1980). Seeing-as, seeing-in, and pictorial representation. *Art and its object* (2nd ed., pp. 205–226). Cambridge: Cambridge University Press.
- Wollheim, R. (1987). *Painting as an art*. Princeton: Princeton University Press.
- Wollheim, R. (1998). On pictorial representation. *Journal of Aesthetics and Art Criticism*, 56, 217–226.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Gabriele Ferretti<sup>1</sup> · Francesco Marchi<sup>2</sup>

Francesco Marchi  
francesco.marchi@rub.de

<sup>1</sup> Eikones - Center for the Theory and History of the Image, University of Basel, Basel, Switzerland

<sup>2</sup> Ruhr-Universität Bochum, Bochum, Germany