

# *Trompe l'oeil* Illusions: Pay (Visual) Attention!

Gabriele Ferretti,\* Francesco Marchi\*\*

\* University of Bergamo

\*\* Ruhr University of Bochum

## Abstract

Considerable interest has been recently devoted to analyzing picture perception and its differences from *vis-à-vis* perception. However, an exhaustive theory of picture perception requires explaining the difference between these two perceptual states and the one we are in when facing pictorial illusions like *trompe l'oeils*, which foster the impression of being in front of a real object available for interaction. One standard story is that these illusions prevent the viewer from perceiving the surface, which is instead possible with usual pictures, this causing the pictorial space to be perceived as real space. In this respect, since recent accounts of pictorial experience revolve around attention, the relation between *trompe l'oeils* and attention should also be accounted for. Nevertheless, nobody has ever offered such a comprehensive explanation. Here, we fill this gap in the literature by offering a theory of *trompe l'oeils* built on the role of attention. This also leads to explain some crucial features of *trompe l'oeil* experience never accounted for within the literature, such as the temporal transiency of *the illusory effect* and the relation between attention and action at its basis. This further clarifies the relation between *picture perception*, *face-to-face perception* and *trompe l'oeil* perception in attentional terms.

*Keywords:* Picture perception, *Trompe l'oeil*, Vision, Action, Attention.

## 1. Introduction

Considerable interest has been recently devoted from both philosophy and vision science to theories aimed at explaining what is the peculiar perceptual state we are in during picture perception, and how it is different from to the one we are in during *vis-à-vis* perception of objects in the flesh.<sup>1</sup>

<sup>1</sup> For a review, see Nanay 2011, 2015, 2016, 2018; Kulvicki 2009; Voltolini 2013; Ferretti 2016, 2017a, 2017b, 2018a, 2018b, 2018c, 2021a; see also Lopes 2005; Hopkins 2012; Matthen 2005.

Notably, such theories also aim at explaining the difference, if any, between these two perceptual states and the one we are in when we are in front of pictorial illusions, like *trompe l'oeils* (Nanay 2015; Ferretti 2020a, 2020b, 2021a). These are a special kind of pictures capable of fostering in the viewer the impression of being in front of an object in the flesh:<sup>2</sup> for example, you may actually be in front of a depicted library, but the illusion deceives you about the presence of a physical library you can effectively interact with (e.g., you can pick up a book) (*Ibid.*).

Recently, there have been parallel but separate attempts at describing the nature of pictorial perception on the one hand (Nanay 2018), as well as at explaining the nature of *trompe l'oeils* illusory power on the other (Ferretti 2020a, 2020b, 2021a; Kulvicki and Nanay 2018; Nanay 2015).

In this respect, on one side, recent approaches aiming at a better explanation of our visual experience of usual (non-illusory) pictures revolves around attention (Nanay 2016, 2018; Ferretti and Marchi 2021).

But, if so, attention would plausibly have a role to play also in our most coherent account of *trompe l'oeil* experience, as well as in determining how we enter these two very different visual experiences. Indeed, an exhaustive theory of picture perception must be capable of explaining *trompe l'oeil illusions* within its framework—a framework that, as said, considers attention.

Nonetheless, on the other side, nobody has ever offered an explanation of *trompe l'oeil* illusory effects by analyzing the mechanisms of attention involved in it.

This leads to a theoretical gap between the literature on ordinary picture perception and the one on pictorial illusions, which would be very beneficial for the literature to be overcome.

Here, we fill this gap in the literature by offering a theory of *trompe l'oeils* that takes into account the role of attention—as said, highly neglected in explanations of *trompe l'oeil* experience and, at the same time, only recently invoked to explain the nature of usual picture perception (*Ibid.*).

However, turning on the alleged role of attention in these illusions will only be the starting point for outlining an account of *trompe l'oeils* that can explain different features of this illusion that have never been accounted for within the literature.

Thus, the present paper offers a novel account of the attentional aspects in *trompe l'oeil* illusions which, *ipso facto*, extends current theories of *trompe l'oeils*. However, this theory is also coherent with recent theories of usual picture perception, and with respect to the standard story on how usual object perception works.

This is very important, for, as said, an explanation of *trompe l'oeils* must be coherent with the explanations of non-illusory picture perception (Nanay 2011, 2015, 2018; Ferretti 2020a; Lopes 2005; Wollheim 1980, 1987, 1998).<sup>3</sup> This is for a simple reason. After all, *trompe l'oeils* are special perceptual objects falling, perceptually speaking, in a sort of middle ground between pictures and ordinary objects: they are pictures, but nonetheless foster in the viewer the impression of an object in the flesh, available to interaction within the physical space. Such an explanation can, thus, settle a complete account of an epistemology of vision concerning these three

<sup>2</sup> For a review, see Nanay 2011, 2015; Kulvicki and Nanay 2018; Voltolini 2013; Ferretti 2016, 2020a, 2020b, 2021a.

<sup>3</sup> Several scholars reflected on the connection between pictures and illusions (Gombrich 1960; Wollheim 1987, 1998; Newall 2010; Bantinaki 2007). However, in this paper, we are interested in the theoretical gap between the literature on ordinary picture perception and the one on pictorial illusions, in the light of the recent debate on pictures in vision science.

perceptual phenomena. To appreciate this point, consider the explanatory benefits of the present account.

1. It considers the role of attention in *trompe l'oeil* visual experience. In this respect, it shows that the role of attention is so important that *trompe l'oeil* visual experience can be massively related to *inattention blindness*.

2. By assessing the role of attention, it offers an explanation for some crucial features of *trompe l'oeil* visual experience, such as its *transiency*, i.e., the fact that the illusion is *time-dependent*. An aspect never accounted for within the literature.

3. In the light of recent idea that action plays a crucial role in generating not only visual experience, but also pictorial experience, with their own differences, as well as in entering *trompe l'oeil* illusions (Ferretti 2021a), the present account suggests an even tighter *relation between attention, vision and action* than what is standardly accepted, in the case of pictorial stimuli.

4. The former three accomplishments lead us to integrate this account within the most recent account on picture perception, which has only focused on the way attention allows us to discriminate between usual picture perception and aesthetic appreciation of pictures, leaving, thus, completely unexplained the *trompe l'oeil* effect in terms of attention (Ferretti and Marchi 2021). This is done by understanding the crucial relation between *picture perception, face-to-face*, and *trompe l'oeil* perception, in attentional terms, in a unified and coherent story.

In what follows, we describe how, in the case of *trompe l'oeils*, (i) the viewer cannot properly exercise visual attention, not even *unconscious* visual attention, upon the surface. This marches in step with the fact that (ii) the pictorial space is built in such a way as to enhance the visual cues responsible for attracting the viewer visual system to the illusion of action possibilities pertaining to the pictorial space, which, however, at that point is perceived as pertaining to a space for interaction, and not as a flat space, related to a bidimensional surface. (iii) This enhancement has, in turn, the role of hiding the flatness of the surface, thus sustaining the illusion (even if temporary) of the possibility of motor interaction with what actually is a pictorial space. We also account for this temporal aspect in terms of attention.

The paper proceeds as follows. We first discuss current views on, respectively, non-illusory picture perception and *trompe l'oeil* perception. Then, after outlining the importance of attention for visual processing, we discuss the role played by attention in picture perception and in generating the illusory effect of *trompe l'oeils*, respectively. Finally, we explain the relation of *trompe l'oeil* to inattentional blindness, the way it relates to action, and its time transiency.

## 2. Usual Picture Perception and *Trompe l'oeil* Perception

Here we briefly describe what we take to be the most up to date view about usual (non-illusory) picture perception and *trompe l'oeil* perception. This will set the ground for appreciating the importance of attention in pictorial experience, and the need for a novel account capable of describing the role of attention also in the case of *trompe l'oeil* visual experience.

Here is the common story about usual picture perception. When perceiving a pictorial object, we visually perceive not just one, but two things: the depicted object, that is the pictorial content, and the picture's surface, that is, the vehicle of the pictorial content. Several arguments have been proposed to suggest that we visually represent them *simultaneously* (for a review, see Nanay 2011; Ferretti 2021a).

However, since visual representations can be either conscious or unconscious, it has been also suggested that we consciously see the depicted object while simultaneously unconsciously seeing the surface—for several reasons we do not need to discuss here, as the recent literature has offered reviews on this.<sup>4</sup> Therefore, according to standard accounts, during picture perception: a. We consciously visually represent the depicted object; b. We unconsciously visually represent the surface.

These two representations occur, as said, simultaneously, when obtaining proper pictorial experience.

Now, with the received view of usual picture perception at hand, we can move onto what has been proposed as a description of what happens in the case of *trompe l'oeil*, on the basis of what we know on usual picture perception.

As we mentioned, a *trompe l'oeil* is a two-dimensional stimulus, a picture, which generates the visual illusion of a three-dimensional space offering interaction. Furthermore, *trompe l'oeil* illusions are *transient* by nature. After a while, due perhaps to subtle movements and shifts in the subjective point of view, the illusion tends to disappear (Nanay 2015; Kulvicki and Nanay 2018; Ferretti 2020a, 2021a, 2023) (we'll get back to this point later in the paper).

Given the account of ordinary picture perception presented above, it has been suggested that the best explanation of the illusory effect of *trompe l'oeil* visual experience is that, in this perceptual case, we cannot perceive the surface, either consciously or unconsciously: our visual system cannot visually track it at all (Ferretti 2020a, 2020b, 2021a).

In this respect, there is an important point on visual simultaneity, usual pictures and *trompe l'oeils*: the strongest philosophical argument supporting the idea that, during usual picture perception, we need to visually represent also the surface along with the perception of the depicted object, is that the best interpretation of experimental results from vision science provides compelling empirical confirmation that, if we do not visually represent the surface, we then enter the illusion of *trompe l'oeil* visual experience, in which the depicted object looks like an object we can motorically interact with. This is because visual *simultaneity* is lost, and we experience a breakdown in pictorial experience (for the detailed philosophical analysis of the experimental results from vision science, see *Ibid.*).

So, with *trompe l'oeil* visual experience—as the explanation goes—the visual representation normally attuned to the surface (in standard pictorial perception), as *per* (b), described in the section above, is missing. Thus, the pictorial content is mistakenly taken, from our visual system, as an object for interaction.<sup>5</sup>

The reader should note that *trompe-l'oeil* visual experience is importantly *repeatable*. On the one hand, the illusion is time-dependent, this meaning that it takes place at a specific time interval, and then expires. On the other, we can be entertained with this illusory effect numerous times, within the same exposure to the painting, in line with the *transient effect* of the illusion. Indeed, we can flip back and forth in between the illusory experience of an object in the flesh and the pictorial experience (Kulvicky and Nanay 2019; Ferretti 2023).

<sup>4</sup> For a review, see Nanay 2011, 2015, 2018; Ferretti 2021a.

<sup>5</sup> The literature in philosophy of perception uses different notions of 'presence' and 'reality' with respect to motor interaction (Matthen, 2005; Ferretti, 2023, 2024, forthcoming a, forthcoming b). Here, we maintain a neutral definition.

In the rest of the article, we offer an explanation, in terms of attention, of the causes of the failure to represent the surface, in terms of attention, that leads to the *trompe l'oeil* illusion, following the intuition above expressed on (b).

### 3. Pictorial Attention

In order to understand the role of attention in *trompe l'oeil* illusion, we need to describe how attention works. There are already compelling accounts available in the literature that examine how attention unfolds in the case of usual ordinary picture perception (Nanay 2011, 2018; Ferretti and Marchi 2021). Here we include a summary of the relevant aspects of this debate that are important for our attempt to adopt and expand such accounts to explain *trompe l'oeil* illusory perception.

There is an old debate about 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 a multiplicity of attentional phenomena, which leads to point to a gradual notion of attention, i.e., a process that has different triggers (endogenous and exogenous), that can be directed at various targets (spatial regions, individual features, objects), and that is located at different stages of cognitive processing (Carrasco 2011; Chun et al. 2011). Most of the attention literature agrees that the function of attention has to do with selecting, prioritizing, and boosting cognitive and perceptual processes. It is also widely accepted that attention has a limited capacity, which depends on the limited metabolic resources of a biological system like the human brain (Kastner and Ungerleider 2001; Lennie 2003), and that external stimuli compete for attentional resources in order to be represented in the visual system (Desimone and Duncan 1995; Duncan 2006). Thus, following established literature on attention and concerning picture perception, it is possible to suggest that:

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.

These two points are supported by empirical considerations about phenomena like *inattention blindness* (Bressan and Pizzighello 2008), *change blindness* (Simons and Rensink 2005), and *hemispatial visual neglect* (Driver and Vuilleumier 2001), in which experiences and reports about presence or changes in visual stimuli are impaired, due to, respectively, cognitive load, perceptual load and neural damage, all of which affect attentional mechanisms.

Back to picture perception, there is an interesting thing to note in the light of this evidence. If both conscious and unconscious visual processing need attention, as this is required for visual processing in general, then, it has been noted that we should modify (a) and (b) accordingly, when describing usual, non-illusory picture perception (Ferretti and Marchi 2021). (a) We consciously visually *attend* to the depicted object; (b) We unconsciously visually *attend* to the surface.

Given the above considerations on visual representations and attention, it follows that a failure in deploying attention to a pictorial two-dimensional surface would entail a lack of representation of said surface, precisely *qua surface*. But, if

that is true, this leads, *ipso facto*, to further hypothesize that the lack of visual representation of the surface that underpins *trompe l'oeil* illusions depends on a temporary failure in the deployment of attention to that surface.

Then, if this is correct, to offer a full explanation of *trompe l'oeils*, we need to understand what causes such a failure in terms of attention. In order to explain this failure, however, we first need to say something on the relation between vision, attention and action.

#### 4. Attention and Action

As we mentioned, during usual picture perception, we always visually represent the surface, even if just *unconsciously*. However, as we discussed above, if unconscious visual representations always require attention as well, as a precondition (more on this below), we also need to unconsciously visually attend to the surface to obtain a proper simultaneity granting a pictorial visual experience.

Therefore, it is also true that, if something could prevent the possibility of having visual attention attuned to the surface, this *would explain* the lack of visual representation of the surface with *trompe l'oeils*, as per our hypothesis.

But why would one think that with *trompe l'oeils* attention to the surface is not in play?

To answer this question, recall that, in the previous section, we mentioned how in the literature on attention it is widely accepted that cognitive and perceptual load can influence attention deployment leading to failures on *noticing, reporting* and *experiencing* stimuli.

In this respect, one of the most compelling illustrations of this characteristic of attentional mechanisms is the phenomenon of *inattentional blindness*.

Inattentional blindness occurs when a subject engaged in an attention-demanding task, which induces load, fails to notice a stimulus that is presented above detection threshold and that is unrelated to the task. The most intriguing demonstration of this phenomenon involves subjects watching a video clip and engaged in counting how many times several players pass a basketball.<sup>6</sup> A significant majority of subjects did not report noticing anything unusual while performing the task, but, in fact, during the video, there is a person dressed in a gorilla suite that suddenly moves into the camera angle, the background of the scene had changed colour, and one of the players was substituted by another. Thus, it seems that if attentional resources are exhausted by a *demanding task*, subjects fail to report significant and clearly visible changes in the perceptual environment.

An important point here is that we hold that attentional mechanisms are involved at all stages of perceptual processing because, due to the limited resources (spatial, temporal and metabolic) available to perception at any given time, perceptual processes need to be selective in their priority of targets, and they need to be efficient and successful in this selection as well. Resource constraints and, hence, also the need for selection, which is achieved by attentional mechanisms, apply to both conscious and unconscious perception.

Now, in the light of what just illustrated, and this is a crucial point, if we had good reasons to think that, in *trompe l'oeil*, attentional resources are also engaged in a *demanding task*, this could deplete them to the point that the system may fail to

<sup>6</sup> Simons and Chabris 1999. A video demonstration can be found at [https://www.youtube.com/watch?v=IGQmdoK\\_ZfY](https://www.youtube.com/watch?v=IGQmdoK_ZfY).

process and visually represent the picture surface, thereby generating the illusion. This would explain why with *trompe l'oeils* attention to the surface is not in play.

But this hypothesis naturally leads to another question: what could then be the analogue of dynamic attention-demanding tasks, like those employed in experiments on inattention blindness, when looking at a static picture such as a *trompe l'oeil*?

We think that the answer to this question lies at the intricate crossroad between *attention*, *perception* and *action*. Let us go more slowly on this.

Perception, action, and attention are crucially linked. Attention directs our perception on the targets for action. Then, action-related cues are of primary source for our visual system. This should clarify in which sense attention directs perception. Attention selects targets for action that are primarily elaborated by perceptual processing, having an ecological significance for the tracking activity of our visual system. In this respect, indeed, for a biological system, tracking visual track of the salient cues for action is even, arguably, a much more fundamental, ecologically valid and demanding task than that, for example, of counting how many times a basketball is passed (cf. above), which is, nonetheless, paradigmatic of the inattention blindness case.

Spotting visual sources of action requires an extreme amount of attentional resources, also and especially given that one of the primary functions of attention is precisely that of selecting, mapping and relating many visual inputs to action outcomes, with the purpose of performing suitable action guidance, as established in the literature (Wu 2011; Watzl 2011a, 2011b).

In this respect, not only does attention guide our actions (Wu 2011; Watzl 2011a, 2011b), but our motor activity can shape the way we exercise attention on the visual scene (Nanay 2021). This is not trivial in the light of the idea that one of the main tasks of vision is to guide action (for a critical overview, see Briscoe and Grush 2015; Nanay 2013; Ferretti and Zipoli Caiani 2021; Ferretti 2021a, 2021b), and that attention directs vision (§3).

Now, we have suggested that (i) with *trompe l'oeils* attention to the surface is not in play, (ii) that this may be due to a demanding task, in this perceptual scenario, and we have also suggested that (iii) vision, attention and action are crucially linked.

In what follows, (iii) will allow us to explain how (ii) occurs, i.e., to explain what is the analogue of dynamic attention-demanding tasks, like those employed in experiments on inattention blindness, when looking at a static picture such as a *trompe l'oeil*, thus satisfyingly confirming (i).

### 5. *Trompe l'oeil*: A Matter of Attention... and Action

How do we explain the attention demanding aspect leading the viewer, with *trompe l'oeils*, not to deploy attentional resources to the surface, this resulting in the illusion? This can be done by taking into account the above-described relation between vision, attention and action. As we shall see, this is reasonable in the light of the fact that, after all, *trompe l'oeil* illusions are *visuomotor illusions*: they deceive about the possibility of vision successfully guiding action toward the stimulus. Following this line of thought, it is straightforward to consider why action-related cues play a special role in our visual economy. These ideas are crucial to understand what happens in the case of *trompe l'oeil*, and with respect to visual attention. Indeed, with these pictorial illusions, the equilibrium between vision, attention and action, correctly reached during usual picture perception, seems to

be broken.<sup>7</sup> Let us now develop our account with these ingredients, by starting from a standard description of these illusions.

In *trompe l'oeil* illusions, the subject typically experiences a pictorial space not as such, but as a space displaying potential possibilities of interaction with one or various objects. When being visually presented with the *trompe l'oeil* of a library, this is apparently displaying the possibility of a grasping action for each of the volumes visually encoded. Even if the picture is static, the system still needs to detect and visually represent a multiplicity of action possibilities, as well as to select which actions to carry out and which not to. After all, also standard, non-pictorial, visual scenes can be static sometimes, and nonetheless foster in us the impression of being available, as they actually are, to motor interaction.

One reasonable description of this perceptual fact is that the detection of action possibilities that are presented within a *trompe l'oeil*, even if only illusorily, seems to be the attention-demanding task that leads to failure to represent the picture surface, *qua surface*, and thus the failure to experience the *trompe l'oeil* as the two-dimensional picture it actually is.

## 6. The Surface and the Illusion

This leads us to an important point. In the case of *trompe l'oeils*, the picture's surface is, by design, rendered a *peripheral feature*, with as little saliency as the artist can manage. This makes hard to visually attend to the surface, *qua surface*. But why is the artist successful in, so to speak, hiding the surface in plain view? In other words, what does she need to enhance or reduce in order for the illusion to take place?

In the light of what previously said on the relation between vision, attention and action, we think that the best explanation for the success of the illusion depends in a large portion on a corresponding enhancement and highlight of illusory action possibilities that are part of the pictorial content, which become flagged by usage of bright colors, contextual cues, and photographic realism (for a very clear example of this, see Vishwanath 2014, and the related analysis, Ferretti 2021a). Let us go more slowly on this.

In every picture, either non-illusory or illusory, there are visual cues related to the pictorial space. Some of these cues are related to action possibilities. But there are also visual cues related to surface. And some of these surface cues are related to action possibilities as well. Take non-illusory pictures. Some cues pertain to the pictorial space. Among those, there are some related to spatial features that, were the depicted object an object in the flesh, would be associated with an action possibility. For example, the specific geometry of a depicted handle of a depicted mug. Of course, these can attract our visual attention. However, in the case of non-illusory pictures, the surface is *easily trackable*, and it is tracked in all its three-dimensionality, as a present object also available for interaction. This is because its action relevant cues can be tracked, and this massively attracts our attention. Thus, in this case, the way the stimulus is built allows the visual system to deploy attention onto the surface, due to its action possibilities.

Then, in usual (non-illusory) picture perception, we need to, at least unconsciously, visually attend to the picture's surface. This leads, in turn, to visually

<sup>7</sup> This is in accordance with recent accounts suggesting that *trompe l'oeils* violate the usual relation between vision and action with pictures (Ferretti 2020a, 2021a).



represent, at least *unconsciously*, the surface *qua surface* as an object offering motor interaction (for example by affording a *visuomotor experience* of flatness as we were to touch it with our hands). Here, again, it is not the flatness of the surface, *per se*, that leads the surface to be tracked and represented, but rather the action possibilities afforded by a flat surface (e.g., stroking instead of grasping), be it a wall, a door, a plank or a canvas. This is what prevents having the impression of possibility for interaction with the pictorial space, which does not look like a space of objects in the flesh (Ferretti 2021a).

When we cannot track the surface, our visual system focuses on the unique available object of perception, i.e., the pictorial space, which, however, is not visually experienced as pictorial, or as depicted, anymore, but rather looks like a space in the flesh offering possibilities for motor interaction (Vishwanath 2014; Ferretti 2020a, 2021a). So, even in usual picture perception, it is the action relevant visual cues that play the pivotal role in engendering a specific type of visual experience.

In accordance with this, what happens in the case of *trompe l'oeils* is slightly different. The picture is built in such a skillful way that the surface is not visible. But how is this achieved? Given what we have previously said about the close relation between attention and visual representation on the one hand, and attention and action on the other, we suggest that the illusion is achieved by manipulation of the visual saliency of the cues related to action possibilities with respect to the surface and the pictorial content. By visual saliency we simply mean the noticeability or capacity of a visual stimulus to attract attention.

In particular, our main claim here is that such a manipulation is achieved by skillful deployment of illumination, shade and depth cues related to the pictorial space's spatial aspects related to an action possibility that, however, are in turn responsible for hiding the spatial features of the surface, which are usually responsible for representing it, exactly *qua surface*, in virtue of related trackable action possibilities. At this point, the spatial features recalling action cues of the pictorial content are more enhanced than those of the surface. As a result, the visual system is attracted by the cues in the pictorial content that are related to spatial features that can be associated with an action possibility. But, at the same time, it neglects those related to the pictorial surface. Then, given (i) the attractiveness of the cues of the pictorial content associated with action possibilities, and (ii) the resulting invisibility of the surface, the visual system is not capable of shifting the attention to the surface (at least at first glance, cf. §7).

To illustrate the differences in the ways in which attention is deployed in the case of an ordinary picture and an illusory one, think about looking at a *natura morta* in a museum. Here, there are immediate and potent cues that attract attention to the surface, precisely *qua surface*. Some examples of these are the frame, the spatial distance between the painting and the wall, the lighting conditions, and the unrealistic proportions of some of the objects depicted. While some of these elements display action possibilities on their own, for example an ornate wooden frame that invites touch and vision to explore it in its own right, some others prevent potential action possibilities on the depicted content to be manifested, for example a light reflection that is not compatible with an object in the flesh in an external (non-pictorial) space. If a skilled artist were to turn such image into a *trompe l'oeil*, they would have to remove these cues by means of eliminating the frame, fixing the lighting conditions and proportions, and so on, to achieve the

result that illusory action possibilities concerning the depicted content emerge as priority targets for attention.

If the artist is successful in hiding all attentional cues that either pertain to the surface, or are incompatible with the depicted content, in case it had been a real object, then the illusion is engendered. As a result, for a time, what is just a picture invites exploration and interaction as if one were indeed in front of a real object. However, as this interaction proceeds, for example by means of a shift of position and perspective on the side of the viewer, the unfolding of the experience violates the (unconsciously formed) expectations of the viewer on how looking at a real object would feel and play out. In other words, this is what invites shifts of attention. This makes the illusion disappear. This suggests that active attentional exploration on these relevant aspects of the piece of art, through time, makes *trompe l'oeils* eludible. Indeed, as noticed, *trompe l'oeils* require that the viewer explores different time intervals, each hosting a change in perspective of the different (illusorily and non-illusorily) cues of the image, in order to flip back and forth with respect to the illusory content of realness and the picture, perceived as such (see Spinicci 2019; see also Ferretti 2020a, 2020b).

This answers all our previous questions. With *trompe l'oeils*, we cannot track the surface. This is because we do not pay attention to the surface. This is because the action related visual cues pertaining to the surface are diminished and there is an attention demanding task concerning the action related cues pertaining to the pictorial content. The presence of enhanced action possibilities related to the pictorial space does not allow us to shift our attention onto the surface, as attention is fully captured by the pictorial content which, at that point, looks like an object offering interaction.

Remember, in this respect, that when the visual system does not enter a simultaneous visual representation of the surface and the pictorial space marked on it, it will consider the only stimulus it receives as an object (arguably, the only external object, in the flesh, that can be tracked). In this case, given the enhanced saliency of the action possibilities of the pictorial space, and the diminished saliency of the surface, the pictorial space becomes the unique object, displaying interaction (*Ibid.*). Since the surface is hidden, and it lurks behind the visual cues related to the action possibilities (of what, however, is just a pictorial space, but does not look at such), enhanced by the artist, which capture our attentional system, these enhanced cues make it difficult to shift our attention onto a surface, given the strong competition between the decreased surface-cues and the enhanced content cues. Thus, we focus on the pictorial space (not perceived, nonetheless, as such), because no surface can be *attentively* tracked. But once our attention is completely captured by the pictorial space, this is because it is fully absorbed by its (illusory) action possibilities. And, in this case, the pictorial space of the *trompe l'oeil* cannot be perceived, indeed, as such anymore. In fact, as said, to enter correct pictorial experience, we need to simultaneously attend to the surface, actually available for interaction, and the depicted object, which is not actually available for interaction (Ferretti 2020a, 2021a). Again, avoiding the possibility of any surface attention leads, accordingly, to a situation in which our attention is fully captured by the unique content of experience: the pictorial space, which, as said, does not look pictorial anymore. So, this is *what* we can attend, given the modulation of attention *where vision-for-action* is attracted from, and *what* it is attracted by.

## 7. A Pictorial (Kind of) Blindness

At this point, the reader should note that, when this happens, the perceptual state is achieving a form of *inattentional blindness to the surface*, which engenders the illusion. This is an important explanatory accomplishment of our account. Accordingly, then, given (i) the artificially lowered saliency of the action relevant cues related to the surface, and (ii) the artificially enhanced visual saliency of the action relevant cues associated with the pictorial content, the visual system is not capable of shifting attention from the pictorial content to the surface. For this reason, it remains attuned to the pictorial content, which however looks, for the reasons exposed, like the content pertaining to an object in the flesh, available to motor interaction.

This move of the visual system seems ecologically reasonable: responding to those visual cues related to action that are more visible. This is perfectly in tune with the idea that, when we cannot track the surface, what is happening is that all the visual resources are devoted to the pictorial space. But here we explain by means of which mechanisms our visual processing is fully devoted onto the pictorial space: the enhanced action possibilities of the pictorial space attract the attentional resources of our visual system, this leading to pay no attention to a surface whose cues are already toned down.

Of course, we can escape the illusion by moving around the picture. What happens here is that different perspectives will balance the visual saliency of the picture's surface with respect to the attentional attraction of the pictorial space. Once we increase surface visibility (usually standing next by one side of it), our attention may be switched and devoted to it, and we may come to perceive the presence of a surface and action possibilities thereof (cf. Ferretti 2020b, 2021a).

If we are right, the skillfully realized design factors of *trompe l'oeils* lead the attentional system to *completely ignore* the presence and flatness of the surface and dedicate all attentional resources to processing and selecting the action-possibilities contained in the pictorial content, which then looks like an object in the flesh.

Note that, in line with what said above, the action possibilities need not be always salient and enhanced in a pronounced manner to obtain the illusion. Even if the image we are confronted with is the one of a depicted umbrella on a wall, instead that the one of a handle, that does not mean that the pictorial content is, always and necessarily, much less attractive to your visuomotor system (though, however, difference of this kind may play, sometimes, a role on the degree of the illusory effect, cf. §9). Action is, most of the time, involved no matter what is visually encoded, and regardless of agent's intentions. And most of the selection of action is done subconsciously (Briscoe and Grush 2015).

If so, our account locates *trompe l'oeil* illusions in the same broad category of phenomena that are due to effects of perceptual and cognitive load on attention. However, we do not claim that these forms of load are the only factor in play during this type of experiences.

## 8. Other (Potential) Illusorily Factors

Of course, other factors may contribute to the artist's success in 'hiding the surface in plain sight'. These may include gestalt elements such as hard-wired perceptual assumptions about luminance sources, where these are skillfully manipulated by the artist to trick the visual system. We also remain pluralist about *trompe l'oeil illusions* themselves, in the sense that we endorse that there may be different types of

experiences that currently fall within the scope of this piece of art and require different types of explanations. We simply contend that our attention-based account offers a neat and widely applicable explanation of how the surface fails to be represented by the visual system during such illusions. Something not offered even by those accounts mentioning lack of surface representation as an explanation for the illusory effect.

Thus, in cases of standard picture perception, the peculiarly enhanced visual spatial cues responsible for obtaining the impression of possibility of motor interaction are missing (with respect to the pictorial space). This is exactly because the surface can be visually tracked and attended to. This is in line with the above-mentioned fact that, with *trompe l'oeils*, our visual system cannot visually represent, even unconsciously, the surface, and thus focuses on the unique object of perception, to which it ascribes all the visual cues available. This is different from the case of standard pictures, which allow the viewer to visually track the surface, at least unconsciously. And this surface tracking is precisely what hinders the possibility of experiencing enhanced visual cues related to the pictorial space, which is indeed perceived as such, i.e., as a bunch of depicted marks on a flat surface.

That said, following our explanation, when we cannot visually track the presence of the surface, this is because we don't have the capacity of exercising visual attention onto the action relevant visual cues afforded by the flat interactable object that is the surface. At this point, we are fooled, at the conscious level, that the depicted object is an object we can interact with, especially if the cues for interactivity with such objects are artificially enhanced: this is precisely what happens in the case of *trompe l'oeil* pictorial illusions, where all our attentional resources are devoted to the action possibilities of the illusory pictorial content.

In this respect, the way we can (or cannot) exercise attention onto the surface, and thus represent it *as a surface*, modulates the way we perceive the depicted object. When our attention cannot track the presence of the surface, that means that our attention is fully captured by what is contained within the pictorial space. And since there is no surface to attend to, the pictorial space does not look pictorial anymore, becoming the unique object of our visual experience. This leads to lack of (even unconscious) surface representation, with consequent unique focus onto the pictorial space (which can be explained also by several optical reasons, see Vishwanath 2014 and Ferretti 2021a).

This is perfectly in line with the recent view that action plays a pivotal role in reaching proper pictorial experience, for when we cannot exercise *vision-for-action* on the surface, *vision-for-action* automatically attunes onto the pictorial space, which is then perceived as an object offering action possibilities. This is why *trompe l'oeils* are so *visually catchy* for our perception of action possibilities (Ferretti 2020a, 2021a).

Now, if this story about the link between action, vision and attention is true, and if this link is in play also during picture perception, as we are claiming here, then, to understand *trompe l'oeil* illusions, not only is it crucial, on the one hand, that *attention directs action* (Wu 2011; Watzl 2011a, 2011b), but also that, on the other, *action can bias our attention* (Nanay 2021). This represents a closure in this theoretical circle, as the crucial role of the relation between attention and *vision-for-action* in both ordinary pictures and *trompe l'oeil* pictorial illusions has been described here.

## 9. A Possible Objection

A final consideration. One may object that a natural explanation for the difference between ordinary pictures and *trompe l'oeils* is that the former includes clear visual

cues to the presence of the picture's surface (such as stereopsis, parallax, reflectance and texture) which are somehow suppressed in the case of *trompe l'oeils*. The attentional attunement to action possibilities would then be, at most, a *consequence* of the success of the suppression of those cues pertaining to the surface.

However, such an explanation is unsatisfactory, as it crucially leaves unspecified how the suppression is achieved. In other words, the crucial point here is that we still need to explain why the cues of the surface, which are still present in *trompe l'oeil* pictures (otherwise we couldn't notice, at some point, the surface) are neglected by the viewer's visual system, in favor of cues pertaining to the pictorial space, such as illusory possibilities for interaction with the depicted object.

Our explanation goes further. It suggests that, granted that (i) the surface is made invisible by the painter, the success of the illusory power of *trompe l'oeils* stems from the fact that, (ii) at first glance, the visual system will be attracted from the spatial cues related to the action possibilities in the pictorial content. Then, since (i) the surface is not visible, and since (ii) action possibilities are tracked onto the pictorial space, (iii) it is hard for the visual system, *prima facie*, to move from the salience of the action possibilities of the pictorial content (not perceived as such) to the invisible surface and its action possibilities (this becomes in principle possible once we embrace a correct angle to spot the flatness of the surface). Note that (i) and (ii) are two crucial aspects in the explanation of the pictorial illusion, *as per* (iii), and explain both (a) how attention works and (b) why it cannot be devoted to the picture's surface, this being not just because the surface is simply invisible, but also because the visual system is attracted by other features that make very hard for it to shift onto another portion of the visual object (i.e., the surface).

Our account has the advantage of offering a unified understanding of both ordinary picture perception and *trompe l'oeil* that goes one step further, proposing a possible and novel *explanans*. The idea is that, in general, the most relevant types of visual cues are, in different circumstances, those pertaining to the action possibilities afforded by the surface and the pictorial content (were it real). This idea, as we have seen, is grounded in standard and widely accepted accounts of the relation between attention, vision and action. Accordingly, both the action relevant cues related to the surface and those related to the pictorial content could be in principle manipulated and exploited.<sup>8</sup> This, in turn, explains how the posited lack of surface representation can be achieved, by such manipulation, in the case of *trompe l'oeil*, with respect to attention.

The take home message is that attention is an enabling condition for surface representation. In the case of *trompe l'oeils*, attention to the motor interaction cues related to the surface is decreased or impaired. Crucially, this is the primary *cause*, rather than just a *consequence*, of the failure to represent the surface as such.

Thus, in a nutshell, our contention is that the sustained failure of representing the surface *qua surface* that engenders a *trompe l'oeil* illusion depends on the enhanced saliency of the possibilities of interaction that would be afforded by the

<sup>8</sup> In this respect, cues to the surface such as stereopsis (i.e., the slight discrepancy between the visual information presented to each eye, which leads to the representation of absolute depth, and to the visual scaling of the object's distance for action; cf. Vishwanath 2014; Ferretti 2018a, 2023) are always, in principle, present in cases of binocular vision (that is, they may be tracked, given the correct exposure to the stimulus), but their mere presence is not sufficient to elicit a representation of a pictorial surface, *qua surface*, unless these specific cues are also attended.

content of the pictorial space were it constituted by real objects, properties and relations (which can be artistically enhanced by various means). Such saliency causes the attentional system to prioritize cues to the pictorial space and neglect cues to the surface, which is, thus, not visually represented. Once action cues are solicited by the pictorial space, they elicit a conscious visual impression of motor interactability, which, arguably, and crucially, *reinforces the need for paying attention to the content of action* (with respect to the pictorial space, which looks like a present space offering reliable action). Thus, the way we can act modulates the way we keep attending to the pictorial content. This is because, once *vision-for-action* is set on a target, its content demands our attention: it's the target to pay attention to for (even if potential) motor purposes. (Indeed, accordingly, motor preparation for action responses is in play even if no intention of acting is at work, and this counts also for standard pictures, cf. Ferretti 2016, 2018b.) This also prevents the possibility of immediately shifting attention to the surface, as soon as we are in front of those illusions. The pictorial space (not perceived as such anymore) is highly attractive for our attentional resources.

Our view has the benefit of making clearly testable empirical predictions. If indeed action possibilities and motor interactability have such a major attention-mediated role to play in *trompe l'oeils*, as we hypothesize, the illusion (i.e., its degree) should then be sensitive to variations in the pictorial content with respects to the (illusory) motor interaction that the content affords. This means that *trompe l'oeil* with contents that vary significantly with respect to motor interaction, e.g., a distant landscape or empty space vs. a cluttered table, should show significant differences in the illusory experience they give rise to, for example in terms of persistence, likelihood of occurrence, vividness, and so on (though even not very enhanced cues may do the basic illusory job, cf. §7).

Furthermore, our account has the clear empirical implication that *trompe l'oeil* illusions should not be achievable with pictorial contents that do not possess possibilities for interaction that are markedly distinct from those of the surface itself, such as depictions of two-dimensional geometrical figures (indeed, usually, the action possibilities pertaining to the surface, and those pertaining to the depicted object, are different; cf. Ferretti 2016, 2018b).

But this is not the end of the story proposed here. On the topic of the *persistence* of the illusion, our account also allows us to explain why the illusion decays after a time. We will explain this in the next section.

## 10. Why the Illusion Is a Matter of Time: *Focal* and *Distributed* Attention

Here we seek to explain the *ephemeral* nature of the *trompe l'oeil* illusion. To do so, we rely on the account of *trompe l'oeil* perception we introduced above, complementing it with the distinction between *focal* and *distributed* attention. Before doing so, a few words on this distinction are of order.

We already discussed, a few pages above, how it is widely accepted that attentional resources are limited. Now, it is also widely acknowledged that these limited resources can be allocated onto the same object, or several objects, in the visual field. In the former case, attention is said to be *focal*. In the latter case, attention is said to be *distributed* (De Brigard and Prinz 2010; Cohen and Dennett 2011; Eriksen and Hoffman 1972; Nanay 2016; Ferretti and Marchi 2021). Importantly, focal and distributed attention can jointly occur during the same visual experience, and flow

seamlessly into one another (Nanay 2016). And they are also in play during usual picture perception: for example, the alternation between these two forms of attention is what determines the difference between a synoptic view of, say, a painting in a museum, where one tries to grasp the overall structure and perspective of the depicted scene, and the zooming in on one tiny detail to enjoy how masterfully the colour was rendered by the artist (Ferretti and Marchi 2021).

So, here is the explanation we offer of the *impermanence* of the *trompe l'oeil* illusions in terms of attentional resources that can be either *focused* or *distributed*.

At the beginning, when the illusion is at its peak, attention is focused onto the depicted object, and distributed *only* across the illusory action possibilities that it affords.<sup>9</sup> When the illusion disappears, the situation changes, and attention becomes distributed between the action possibilities afforded by the surface and the visual properties of the depicted object, which are not action relevant anymore, like in ordinary picture perception. In this case, only the action possibilities of the surface are available.<sup>10</sup> In other words, in the first case, attention, distributed or focal, is completely attuned to the pictorial space. In the second case, attention is divided in between the pictorial object and the surface, as with the simultaneity of usual picture perception.

Let us now offer some more details that are crucial for this explanation. One common sensorimotor aspect of *usual* (non-illusory) picture perception is that we do not perceive spatial shifts with respect to the pictorial content, but just concerning the surface: the *Mona Lisa* will keep looking at you as you move with respect to the painting (Hopkins 2012; Nanay 2010, 2011, 2015, 2018; Lopes 2005; Ferretti 2020b, 2021a). Note that the spatial shifts are not illusory, as these are something we should perceive but that, for visual computational reasons, we can't (Vishwanath et al. 2005; Ferretti 2020b, 2021a).

With *trompe l'oeils*, however, given the violation of usual perception in the pictorial space, it is possible to correctly perceive the *actual* sensorimotor shifts with respect to the pictorial content, as the surface is, initially, missing from perception (Ferretti 2020b).

Bear in mind that, with the *trompe l'oeil* illusions, after a while, the illusion disappears, and what looks like an object for interaction comes back, to our sight, looking like something embedded within the pictorial space.

This is, arguably, due to the fact that the perceived sensorimotor shifts generate expectations of how the scene would unfold were the depicted object indeed in the flesh (Vishwanath et al. 2005; Nanay 2010; Ferretti 2020b, 2021a).

Above, we argued that the illusory possibilities for interaction (allegedly) offered by the depicted object exhaust attentional resources, which then fail to be allocated to the representation of the surface. This is how the illusion arises. While not performing big movements, the spatial shifts of the *trompe l'oeil* are detected, and they apparently pertain to a real object, in line with what said a few lines above (as the surface cannot be tracked). After some time, however, when the subject moves in a more enhanced manner, and perceptually explores the scene actively, the expectations about sensorimotor shifts are violated, as the object is

<sup>9</sup> Nanay (2016) makes a similar point about aesthetic attention that is focused on the content and distributed across its properties (cf. Ferretti and Marchi 2021).

<sup>10</sup> But, compared to the action possibilities originally fostered by the pictorial content (e.g., the multiple grasping acts on an umbrella) the one of the surface is quite simpler: we may simply realize we can touch it.

only depicted and not in the flesh (this is perfectly in tune with the accounts explaining sensorimotor shifts in picture perception, cf. above). The more the subject moves, the more change in perspective allows the surface to come into play in perception. At that point, the perceived object will not offer any spatial, sensorimotor change, as pictorial objects can't display spatial shifts (cf. above)—while *trompe l'oeils* can. Such violations accrue and eventually become salient enough as to take attention away from the object and, thus, to bring it on themselves. The problem is then solved by the visual system by rearranging the deployment of attention in a way that is now (correctly) sensitive to the (presence of a) surface (to its action possibilities), as this is the only interpretation of the visual scene that is consistent with the actual changes in sensorimotor shifts (in accordance with the last account of sensorimotor shifts in both pictures and *trompe l'oeils*, cf. Ferretti 2020b, 2021a).

To summarize this very last point, when first seeing the *trompe l'oeil* picture, the system generates expectations about sensorimotor shifts relative to an apparently present object, shifts which can be visuo-spatially derived. These expectations are soon violated, and sensorimotor experience becomes consistent with the one a depicted object is supposed to offer. Such violations, then, accrue to the point that they become salient enough to take attention back from the illusory action possibilities afforded by the pictorial content, and deploy them on those of the surface as well, turning *trompe l'oeil* perception back into usual picture perception.

While at the beginning sensorimotor shifts are perceived as if belonging to a real object, now they cannot be appreciated, with respect to what is a pictorial content.

This dynamic is clearly allowed by our account of the illusion, and explains the temporal aspect and the *ephemeral*, transient temporal nature of *trompe l'oeil* perception.

## 11. Some Clarifications on Our Claim

There are some important points to consider here, regarding the novelty of our claim, and its relation to similar cases.

### 11.1 Novelty

Let us start with the novelty. First, there is no paper focusing on the relation between *trompe l'oeils* and attention, so much that accounts working on picture perception and attention suggested that, if also *unconscious* attention exists, then, we should try to explain *trompe l'oeils* by referring to what happens to attentional mechanisms in this case of pictorial illusion (Ferretti and Marchi 2021). But this was left open in the literature, and as a further step forward, not embraced in these accounts. Second, *a fortiori*, the claim that *trompe l'oeils* foster the impression of an object in the flesh because of the attentional resources lacking on the surface has never been defended, nor explicitly mentioned, with an argument (*a fortiori*, as the most important explanation), within the literature (with *trompe l'oeils*, by definition, simply being pictures that deceive the spectator into the illusory experience of an object in the flesh). Third, *as per* the explanation of this phenomenon, scholars agree that something is happening at the surface level, when trying to explain what constitutes the illusion. But whether this is because of (1) unconscious perception of the surface or (2) lack of any surface tracking (conscious or unconscious), is what was at the heart



of the debate within the literature (for a review, see Ferretti 2020a, 2021a). That said, whether (3) attention plays some role in this illusion was not at all considered within the literature (indeed, there is no explicit distinction between conscious and attentive states, see Ferretti 2020a: 35 and footnote 11). In this respect, it has been suggested that option (2) provides the best explanation (cf. Ferretti 2020a, 2020b, 2021a). And we complement this explanation with attention (pun intended). To conclude, the merit of this paper is precisely that of offering an account of the illusory effect of *trompe l'oeils* by describing the link between perception, attention and action (in relation to their conscious and unconscious counterparts).

### 11.2 Action, again

However, it is also worth mentioning other cases, interesting for the study of *trompe l'oeils*, somewhat related to action effects. In this respect, it is worth explicitly nothing, following (§ 10), that *trompe l'oeils* also offer *anamorphic effects* (Ferretti 2020b). And, *as per* these effects, the painting (illusorily perceived as a real object) shows the perspectival experience of an object in the flesh, and displays changes in the visual content as the viewer moves with respect to it, contrary to the fact that, in standard non-illusory pictures, the picture always looks the same, and no perspectival aspect is noted. This effect is also related to action, and indeed on the sensorimotor understanding of the different perspectives changing as the point of view changes (for a review of this point, see Ferretti 2020b, 2021a). In this case, as the surface is not perceptually tracked, sensorimotor understanding is deployed toward the pictorial space, which offers *anamorphosis*. In standard cases, perceptual tracking of the surface, which of course also depends on attention (see Ferretti and Marchi 2021), is what permits to have sensorimotor understanding of the surface, this hindering sensorimotor understanding with respect to the depicted object. Importantly, standard descriptions of *trompe l'oeils*, which also take into account *anamorphosis*, do not consider, however, any role of attention (Ferretti 2020b, 2021a).

That said, the biggest effect in *trompe l'oeils* is that of the illusion of interaction, so much that *anamorphosis* frequently takes, perceptually, a back seat (Ferretti 2021a). For this reason, our account aims at explaining the (a) illusion of interaction by taking a look at how attention plays a role within this illusion. However, it is easy to infer that lack of attentional resources upon the surface is also what is at the basis of (b) lack of sensorimotor understanding with respect to the surface. No surface is tracked. Indeed, no attention is deployed toward the surface. Thus, there cannot be any sensorimotor understanding with respect to the surface.

All in all, this explains why our point here is crucial to understand the role of perception and action in *trompe l'oeils*, by adding the contribution of attention, to perception and action, in the light of the distinction between conscious and unconscious attention. This, with respect to the different ways vision and action, and thus attention, can interlock: (i) vision guides action, (ii) vision is a form of sensorimotor process, and (iii) action can bias vision (cf. Briscoe and Grush 2015; Ferretti and Zipoli Caiani 2021).

### 11.3 Hyperrealism and *Trompe l'oeils*

Another important point is about the distinction between *trompe l'oeil* and *hyperrealist* pictures, and experiences thereof. These two perceptual scenarios are different. Hyperrealist (though not illusory) paintings foster, for sure, a very vivid visual

experience of the pictorial space, and effectively more vivid than the one we obtain with standard pictures. But, as with standard pictures, the surface is visible: vivid visual experience of the pictorial space is not as enhanced as the one we can obtain from real objects, and this counts also for the real object that the surface is. The pictorial visual features cannot be compared within these two cases. Thus, though more enhanced in the visual experience of the pictorial space, hyperrealist pictures are not illusory, as they do not deceive the eye (i.e., *trompe l'oeil*) (see Ferretti 2018b, especially footnote 30).

#### 11.4 Usual and Illusory Perception of Pictures

One further interesting point is the following. With *trompe l'oeils*, visual attention in particular, and visual perception in general, are totally captured by the depicted object, as a result of surface invisibility. We may wonder about whether this is just what happens in standard cases of depiction, though to a lesser degree. The reader should note that, while it is reasonable to say, in folk terms, that most of the time we are in front of a picture we pay attention to the depicted object (Nanay 2011), while ignoring the surface, it remains true that, technically speaking, we are always exercising our attention toward both the surface and the depicted objects, though this attention can behave differently (Ferretti and Marchi 2021). The case of *trompe l'oeil*, thus, wherein the surface can't, *prima facie*, be tracked, is not just a weaker case of picture perception: indeed, pictorial experience (or *pictoriality*) is lost (during the illusion, of course) (Ferretti 2021a). In this case, the difference is not in quantity of attention, or degree, but in the quality of the processing and the related experience, given by the different ways in which the visual system can process the object it is in front of. Again, being in front of a surface does not entail we can, automatically, visually represent it (Nanay 2011, 2015, 2016; Ferretti 2021a). In *trompe l'oeils*, indeed, we don't represent it, at least at the beginning, when the illusion takes place. This is hardly comparable to the fact that, in ordinary, non-illusory perception, we focus more on the depicted object (as the surface is indeed visible, and, *qua* visible, also tracked, at least unconsciously). Surface processing (its correct tracking) is what makes these two perceptual scenarios different.

#### 11.5 The Painter and the Perceiver

A final interesting point, related to the previous one, is the following. Given that the failure to visually track the surface can be typically overcome through time, *trompe l'oeils* may simply be taken as paradigmatic pictures upon which the deployment of attentional resources is made harder by the painter, but in the end succeeds as in standard experiences of depiction.

For sure, the deployment of attentional resources is made harder by the painter. But, as explained concerning the former point, this problem with attention makes *trompe l'oeils* and standard pictures different, at least when the illusion takes place. Of course, it is true that, after the illusion has been spotted, the experience is that of a pictorial object. That said, it is important to consider that the illusion can be intermittent: after I escaped the illusion, I can always be deceived subsequently. At time  $t_1$  I am deceived. At time  $t_2$  I can escape the illusion. At time  $t_3$  I can be deceived again (cf. Ferretti 2020a, 2021a, 2023).

To conclude, on our view, *trompe l'oeils* are much more complex visual objects than usually considered, which can host, at different time intervals, the

illusion of an object in the flesh, as well as the correct perception of a pictorial object. Thus, they are multi-layered objects of perception, whether pictorial or not, depending on time.

## 12. Conclusion

Standard accounts of picture perception tried to explain *trompe l'oeil* illusions by pointing out the difference between usual (non-illusory) picture perception, *trompe l'oeil* perception, and ordinary perception of concrete objects (Nanay 2015; Ferretti 2021b).

But only recently, our best theory of picture perception has invoked the importance of attention to explain what's so special about the visual state we are in during picture perception (Nanay 2011, 2016, 2018; Ferretti and Marchi 2021).

However, a coherent theory of *trompe l'oeil* illusory experience must fit this new story about pictorial attention (Ferretti and Marchi 2021). But nobody has ever tried to upgrade the current theories of *trompe l'oeil* as to involve an explanation relying on the mechanisms of visual attention.

We filled this gap by explaining the *visuomotor illusion* fostered by *trompe l'oeils*, by invoking the deep links between attention, vision and action.

This is crucial for theoretical coherence toward (a) a theory of object perception, (b) hosting a theory of picture perception, (c) offering a theory of pictorial illusions, (d) which explains these middle way perceptual phenomena *trompe l'oeil* visual experiences are. And this is also crucial for offering a more compelling story about the visual mechanisms leading to, or not to, ascribe the possibility of interaction to a pictorial object, while explaining how standard pictorial experience is lost if these mechanisms are not properly at work.

Finally, this further expands our understanding of the crucial relations between attention, vision and action, in the light of the recent and controversial notion that action possibilities<sup>11</sup> play a pivotal role in picture perception, explaining how we may fall into a *trompe l'oeil* illusions (Ferretti 2021a).<sup>12</sup>

## References

- Bantinaki, K., 2007. Pictorial perception as illusion. *British Journal of Aesthetics*, 47 (3), 268–279. <https://doi.org/10.1093/aesthj/aym007>

<sup>11</sup> The shrewd reader may wonder about why we do not use the notion of affordance (Gibson 1979) when talking about action possibilities. This is for some simple reasons. First, the literature on pictures and action does not refer to this notion, but only to the notion of action possibility or action property (Ferretti 2016, 2019, 2021b; Ferretti and Zipoli Caiani 2024). Second, the notion of affordance usually finds its place in a debate about what is the best account of the mind, whether as an ecological non-representational system or as a representational system. We don't want to step into this debate (Chemero 2009). Of course, the notion of affordance successfully captures the notion of an action possibility afforded, indeed, by a visual target. That said, behind this notion also lie numerous theoretical expectations we would prefer not to embark on.

<sup>12</sup> Both authors contributed equally to the arguments and the writing of this article. The order of authors is purely alphabetical. Gabriele Ferretti would like to acknowledge support from a Humboldt Fellowship, hosted by Professor Albert Newen at the Institute for Philosophy II, Ruhr-University Bochum, Germany. Both authors would like to thank Albert Newen and Bence Nanay for offering comments on previous versions of the paper.

- Bressan, P. and Pizzighello, S. 2008. The attentional cost of inattentive blindness. *Cognition*, 106 (1), 370–383. <https://doi.org/10.1016/j.cognition.2007.03.001>
- Briscoe, R., and Grush, R. 2015. Action-based theories of perception. In: E.N. Zalta, ed., *The Stanford encyclopedia of philosophy*. <https://www.plato.stanford.edu/archives/fall2015/entries/action-perception/>
- Carrasco, M. 2011. Visual attention: the past 25 years. *Vision Research*, 51 (13), 1484–1525.
- Chemero, A. 2009. *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Chun, M.M., Golomb, J.D., and Turk-Browne, N.B. 2011. A taxonomy of external and internal attention. *Annual Review of Psychology*, 62 (1), 73–101.
- Cohen, M.A. and Dennett, D.C. 2011. Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, 15 (8), 358–364.
- De Brigard, F. and Prinz, J. 2010. Attention and consciousness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1 (1), 51–59.
- Desimone, R. and Duncan, J. 1995. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18 (1), 193–222.
- Driver, J. and Vuilleumier, P. 2001. Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition*, 79 (1-2), 39–88.
- Duncan, J. 2006. EPS Mid-Career Award 2004: brain mechanisms of attention. *Quarterly Journal of Experimental Psychology (Hove)*, 59 (1), 2–27.
- Eriksen, C.W.W. and Hoffman, J.E. 1972. Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics*, 12, 201–204.
- Ferretti, G. 2016. 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. 2017a. 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. 2017b. Pictures, emotions, and the dorsal/ventral account of picture perception. *Review of Philosophy and Psychology*, 8 (3), 595–616. <https://doi.org/10.1007/s13164-017-0330-y>
- Ferretti, G. 2018a. Visual feeling of presence. *Pacific Philosophical Quarterly*, 99, 112–136. <https://doi.org/10.1111/papq.12170>
- Ferretti, G. 2018b. The neural dynamics of seeing-in. *Erkenntnis*, 84 (6), 1285–1324. <https://doi.org/10.1007/s10670-018-0060-2>
- Ferretti, G. 2018c. The nature of pictorial representations. *Phenomenology and Mind*, 14, pp. 136–144. [https://doi.org/10.13128/Phe\\_Mi-23631](https://doi.org/10.13128/Phe_Mi-23631)
- Ferretti, G. 2019. Visual phenomenology versus visuomotor imagery: How can we be aware of action properties? *Synthese*. <https://doi.org/10.1007/s11229-019-02282-x>
- Ferretti, G. 2020a. Why *trompe l'oeils* deceive our visual experience. *The Journal of Aesthetics and Art Criticism*, 78 (1), 33–42. <https://doi.org/10.1111/jaac.12688>
- 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. <https://doi.org/10.1111/jaac.12750>
- Ferretti, G. 2021a. Why the pictorial needs the motoric. *Erkenntnis*, 88, 771–805. <https://doi.org/10.1007/s10670-021-00381-1>

- Ferretti, G. 2021b. A distinction concerning vision-for-action and affordance perception. *Consciousness and Cognition*, 87. <https://doi.org/10.1016/j.concog.2020.103028>
- Ferretti, G. 2023. For an epistemology of stereopsis. *Review of Philosophy and Psychology*. <https://doi.org/10.1007/s13164-023-00711-y>
- Ferretti, G. 2024. Understanding reality and presence in dreams through imagery. *Analysis*, <https://doi.org/10.1093/analys/anae025>
- Ferretti, G. *forthcoming*, a. Philosophical implications of derealization disorder. *Synthese*.
- Ferretti, G. *forthcoming*, b. What stereoblindness teaches us about visual reality. *Phenomenology and the Cognitive Sciences*.
- Ferretti, G. and Marchi, F. 2021. Visual attention in pictorial perception. *Synthese*, 199, 2077-2101. <https://doi.org/10.1007/s11229-020-02873-z>
- Ferretti, G. and Zipoli Caiani, S. 2021. Between vision and action: Introduction to the special issue. *Synthese*, 198, 3899–3911. <https://doi.org/10.1007/s11229-019-02518-w>
- Ferretti, G. and Zipoli Caiani, S. 2024. An all-purpose framework for affordances. Reconciling the behavioral and the neuroscientific stories. *Synthese, Special Issue: Neuroscience and Its Philosophy*. <https://doi.org/10.1007/s11229-024-04646-4>
- Gibson, J.J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gombrich, E. 1960. *Art and illusion*. New York: Pantheon.
- Hopkins, R. 2012. Seeing-in and seeming to see. *Analysis*, 72, 650–659.
- Kastner, S. and Ungerleider, L.G. 2001. The neural basis of biased competition in human visual cortex. *Neuropsychologia*, 39 (12), 1263–1276.
- Kulvicki, J. 2009. Heavenly sight and the nature of seeing-in. *Journal of Aesthetics and Art Criticism*, 67 (4), 387–397.
- Kulvicki, J. and Nanay, B. 2018. Art made for pictures. *Phenomenology and Mind*, 14, 120–134.
- Lopes, D.M. 2005. *Sight and sensibility. Evaluating pictures*. Oxford: Oxford University Press.
- Lennie, P. 2003. The cost of cortical computation. *Current Biology*, 13 (6), 493–497.
- Matthen, M. 2005. *Seeing, doing and knowing: a philosophical theory of sense perception*. Oxford: Oxford University Press.
- Nanay, B. 2010. Transparency and sensorimotor contingencies: do we see through photographs? *Pacific Philosophical Quarterly*, 91, 463–480.
- Nanay, B. 2011. Perceiving pictures. *Phenomenology and the Cognitive Sciences*, 10, 461–480.
- Nanay, B. 2013. *Between perception and action*. Oxford: Oxford University Press.
- Nanay, B. 2015. *Trompe l'oeil* and the dorsal/ventral account of picture perception. *Review of Philosophy and Psychology*, 6, 181–197.
- Nanay, B. 2016. *Aesthetics as philosophy of perception*. Oxford: Oxford University Press.
- Nanay, B. 2018. Threefoldness. *Philosophical Studies*, 175, 163–182. <https://doi.org/10.1007/s11098-017-0860-2>
- Nanay, B. 2021. Perception is not all-purpose. *Synthese*, 198, 4069–4080. <https://doi.org/10.1007/s11229-018-01937-5>
- Newall, M. 2010. *What is a picture? Depiction, realism, abstraction*. London: Palgrave Macmillan.

- Simons, D.J. and Rensink, R.A. 2005. Change blindness: past, present, and future. *Trends in Cognitive Sciences*, 9 (1), 16–20.
- Spinicci, P. 2019. Temporal images: The anamorphic game and the nature of picture. In: J. Pelletier and A. Voltolini, eds. *The pleasure of pictures, pictorial experience and aesthetic appreciation*, New York: Routledge, 162–181.
- Vishwanath, D., Girshick, A.R., and Banks, M.S. 2005. Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, 8, 1401–1410.
- Vishwanath, D. 2014. Toward a new theory of stereopsis. *Psychological Review*, 121 (2), 151–178.
- Voltolini, A. 2013. Why, as responsible for figurativity, seeing-in can only be inflected seeing-in. *Phenomenology and the Cognitive Sciences*, 14 (3), 651–667.
- Watzl, S. 2011a. The nature of attention. *Philosophy Compass*, 6 (11), 842–853.
- Watzl, S. 2011b. The philosophical significance of attention. *Philosophy Compass*, 6 (10), 722–733.
- Wollheim, R. 1980. Seeing-as, seeing-in, and pictorial representation. In *Art and its object*. Cambridge: Cambridge University Press, 205–226.
- 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.
- Wu, W. 2011. Attention as selection for action. In: C. Mole, D. Smithies, and W. Wu, eds., *Attention: Philosophical and psychological essays*. Oxford: Oxford University Press, 97–116.