

# Seeing-In as Aspect Perception

Fabian Dorsch

In: Gary Kemp & Gabriele Mras (eds.), *Wollheim, Wittgenstein, and Pictorial Representation*, London: Routledge, 2016.

1. In this chapter, I argue that seeing-in, the central element of pictorial experience, is a form of aspect perception. The argument thus connects Wollheim's main contribution to the philosophy of depiction with one of the central themes in Wittgenstein's philosophy of psychology.<sup>1</sup> This *Aspect View* of pictorial experience that I would like to put forward is designed to be a direct competitor to the *Experienced Resemblance View* and the *Imagination View*, and to improve on both of them by incorporating some important elements of either. More specifically, the Aspect View claims that seeing-in involves the imperfect illusion of the picture's surface as possessing the aspect of having the visual appearance of a three-dimensional arrangement of objects (i.e. the depicted scene). And, as part of this aspect perception, we both experience the picture's surface as resembling the depicted scene in two-dimensional shape and have a non-perceptual awareness of the depth and volume of that scene which is similar to, but not quite like imagining.

The argument proceeds in five steps. The first (§§ 2-7) is primarily concerned with discussing certain important features of pictorial experiences, including their relation to pictures. I start with distinguishing two important aspects of pictorial experience, namely that it makes us visually aware both of a surface's status as a picture and of the latter's specific pictorial content. Then, after introducing Wollheim's idea that these two properties of pictures (i.e. their status and their content) are response-dependent specifically on experiences of seeing-in, I spell out in more detail what seeing-in is like, with a special focus on its twofoldness. Following Wollheim in identifying the central element in recognising something as a picture with the visual awareness of the depth of the depicted scene, I conclude the preparatory first step of the argument by raising two crucial questions — the Question of Content about how we recognise that a given picture depicts *X* rather than *Y*; and the Question of Depth about how we recognise the depth of the depicted scene.

The second step of the argument (§§ 8-12) establishes the distinction between divisive and unitary accounts of seeing-in and finds fault with Wollheim's endorsement of a divisive view, as well as with his conception of the precise difference between the two folds of seeing-in, the configurational and the recognitional fold. The third (§§ 13-16) and the fourth step (§§ 17-21) each critically discuss one of the two most

---

<sup>1</sup> The idea that pictorial experience is an instance of seeing an aspect can also already be found in Wittgenstein's writings, albeit not yet combined with an understanding of pictorial experience in terms of seeing-in in Wollheim's technical sense (see, e.g., Wittgenstein 1953/2009 part II § xi and Wittgenstein 1958 pp. 163ff. , as well as Budd 1987 and Budd 1991 ch. IV for discussion.

prominent unitary accounts of perceptual experience, the Experienced Resemblance View and the Imagination View, respectively. While the main difficulty for the first view is that it has no resources to answer the Question of Depth, the second view fails to provide a proper reply to the Question of Content and, moreover, misidentifies our awareness of what is depicted as imaginative. As a viable alternative to both views, I present and begin to defend the Aspect View in the fifth and last step of the argument (§§ 22-26), not the least by spelling out in a bit more detail what aspect perception generally involves, and why seeing-in should be understood as an instance of it. In particular, the Aspect View is in a position to answer both the Question of Content and the Question of Depth.

2. A *pictorial experience* is one in which we are *visually* aware of a flat surface *as a picture*. This means that, whenever one of our pictorial experiences is veridical, we succeed in visually recognising some picture for what it really is, namely a two-dimensional depiction of some objects and their visible features – and not just a flat surface.<sup>2</sup> In addition, merely judging and coming to know that something is a depiction of something else is not an instance of pictorial experience because we are not aware of the object’s status as a picture in a visual manner.

3. Visually experiencing a picture as a picture necessarily involves at least two aspects of awareness. First, we have to be visually aware of the experienced object’s property of being a picture. That is, we should visually experience it as being a picture, rather than merely as being a flat surface, say. Second, we have to be visually aware of its pictorial content, of the specific objects and features that it happens to depict. That is, we should have a visual experience of *X* rather than *Y*, say, if the picture depicts *X* rather than *Y*. In other words, to visually recognise something as a picture means to visually recognise both *that* it is a depiction and *what* it is a depiction of. It is impossible to enjoy a pictorial experience, in the sense specified, that includes only one of the two elements.

On the one hand, we cannot experience something as a picture without being aware of (at least part of) its pictorial content. For the only perceivable feature in virtue of which I can visually recognise something as a picture is precisely its possession of a pictorial content. Pictures do not show any other perceivable characteristics which indicate their pictorial nature. For instance, just being aware of a flat surface with marks of colour on it does not suffice for experiencing the surface as depicting something. We also have to be aware, at least to some extent, of what it depicts. Therefore, although it is not required that one has to be aware of the

---

<sup>2</sup> My notion of pictorial experience (and seeing-in) is thus narrower as Lopes (2005)’s and Newall (2009)’s notions which are also meant to apply to our standard experiences of illusionist and actualist pictures (i.e. experiences which are introspectively indistinguishable from potential face-to-face perceptions of the three- and two-dimensional objects that those pictures depict). Besides, I concentrate in what follows exclusively on two-dimensional pictures. See Hopkins 2004, Hopkins 2010b and Martin 2012 for discussions of whether there are also three-dimensional pictures.

whole pictorial content of an object in order to recognise it as a depiction, awareness of something as a picture implies the simultaneous awareness of (part of) what is depicted. The same applies to the specificity of pictorial content. While we might not always succeed in recognising the particular nature of the depicted object when having a pictorial experience, such an experience always involves the recognition of at least some determinable aspect of its nature — minimally, that is an object or figure distinct from the depicting surface.

On the other hand, as already hinted at, if we are merely aware of what a picture depicts, but do not recognise that it is indeed a picture, then we end up having a (potentially misleading) perceptual experience, rather than a pictorial experience. There are two possibilities here, depending on whether the picture concerned depicts three- or two-dimensional objects. Trompe-l'oeils belong to the first kind of pictures. When we look at them and fail to experience them as pictures, it (wrongly) seems to us as if we really see a three-dimensional scene, while this scene is in fact merely depicted. Good examples of pictures of the second kind are actualist pictures, that is, pictures which depict flat surfaces that look exactly like the pictures themselves (Jasper Johns' *Target Paintings* are representative cases). Again, when we look at such pictures, we simply see nothing but the visual appearance of the two-dimensional object that is depicted (which also happens to be the visual appearance of the picture itself). In fact, in contrast to trompe-l'oeils which we can normally recognise as such just by changing our point of view on them, we can arguably never visually recognise actualist pictures as pictures, given that their visual appearance is identical with, and hence cannot be visually discriminated from, the visual appearance of what is depicted.

4. None the less, the two aspects of pictorial experience concern two distinct facts about the nature of depictions, which show a certain kind of independence of each other. First of all, that something is a picture has no influence on what it depicts. This suggests that those properties of a picture that determine its status as a depiction are, taken on their own, not sufficient for the determination of its pictorial content. It is difficult to see how, otherwise, it could be explained that pictures with very different subject matters — and with very different degrees of specificity and complexity of pictorial content — still count as instances of the same kind (i.e., as depictions). The best explanation seems to be that some shared feature determines their status as pictures, but not — or not on its own — their pictorial content.

But it also seems true that what specifies the pictorial content is, at least to some extent, independent of what determines the status as a depiction. In particular, we can easily change what a picture depicts (e.g. by applying more paint), without influencing its property of being a picture. So there appear to be features of pictures that are central to the determination of what they depict, but not of their status as depictions. The idea is thus that we can distinguish two sets of properties of a picture, one of which is responsible for their being a picture, while the other

determines what it depicts; and that the two sets of properties are to some degree independent of each other, even though they presumably overlap and cannot be instantiated completely independently of each other, given that nothing is a picture without a minimally determinate pictorial content, and vice versa.

5. One of Wollheim's important contributions to the philosophy of depiction was his emphasis of the fact — already noted by Gombrich (1960) — that the properties in question are response-dependent (though he did not put it this way). Both that something is a picture and what it depicts depend on how we normally respond to it.<sup>3</sup> More specifically, something is a picture of something else only if, and partly because, we normally visually recognise it to be a picture of that other object (assuming that we are looking at the picture). If we act on our intention to produce a depiction of a certain object, but the resulting marks on the surface concerned are such that no one (us included) can become visually aware of the surface as a picture, or identify what it is supposed to depict, then it is not a depiction of the object in question, and we have failed in our attempt to produce one.

In other words, both being a picture and having a certain pictorial content are not strongly objective properties in the sense that they could be instantiated in the world without any (human) beings ever possessing the capacity to visually recognise things as pictures of something else. Instead, they are objective properties only in the sense that they allow for intersubjectivity and misrecognition.<sup>4</sup> But the two properties are also not strongly subjective in the sense that, whenever we visually experience something as a picture of something else, then it is such a picture. For we may err about whether something is a picture, as well as about what it specifically depicts if it is a picture. This is why what matters for the presence of depiction are our normal responses, which happen under conditions that are suitable for veridical pictorial experience. It might not be easy to characterise these conditions, but they are likely to be very similar to those for visual perception.

Of course, like most, if not all response-dependent properties, being a depiction and depicting something specific are features that are partly grounded in the lower-level properties of the pictures concerned. Given that we can modify what a picture shows us by applying or removing paint, the colour, texture and shape of marks of paint plays some role, yet to be specified, in the determination of what is depicted. Similarly, some aspect of the marks on the surface should also be relevant for whether it is a depiction in the first place.

6. Now, the fact that the properties of being a depiction and of having a certain

---

<sup>3</sup> According to Wollheim, at least the status as a depiction also depends on the artist's intention to produce a picture. Clouds in the sky or marks on a wall do not depict something, even if we can 'see' something 'in' them, because they have not been intended to give rise to such an experience (Wollheim 1987 pp. 47ff.).

<sup>4</sup> See McDowell (1984/1998) for this distinction between two senses of objectivity.

pictorial content are response-dependent raises the question about the nature of the normal response(s) concerned. Given that, as just argued, we cannot be aware of one of the properties without being aware of the other, the response by means of which we recognise a thing as a picture and the response by means of which we recognise what specifically this picture depicts should be one and the same. This is indeed what Wollheim proposes. The core of his view is the claim that we visually recognise a flat surface as a picture of some object or scene by means of ‘seeing’ the object or scene ‘in’ the picture. This visual experience of *seeing-in* is, minimally, characterised by five features (Wollheim 1986 §§ 2f. Wollheim 1987 pp. 46f. and p. 62; Wollheim 1998 p. 221).

First, it involves the visual awareness of the flat surface and its visible features, including its flatness and the colour, shape and arrangement of the marks on it (i.e. what Wollheim calls the ‘configurational fold’ of seeing-in). Second, seeing-in involves the visual awareness of what is depicted, that is, of a certain spatial arrangement of two- or three-dimensional objects and their visible features (i.e. what Wollheim calls the ‘recognitional fold’ of seeing-in). Third, our visual awareness is always distributed between the configurational and the recognitional fold, meaning that we are never only aware of one of them. In other words, we are always aware both of the depicting surface and of what it depicts. This third feature is what Wollheim calls the ‘twofoldness’ of seeing-in.<sup>5</sup>

Fourth, the two folds of awareness do not constitute two independent experiences that merely happen to occur simultaneously, but form a single, unified experience. That is, the configurational and the recognitional fold could not exist in isolation from each other, or at least not in exactly the same experiential way in which they form part of seeing-in (e.g. they could not have the same phenomenology if they were to occur on their own; Hopkins 1998 pp. 20f.). After all, there is an experiential difference between seeing a person face-to-face and seeing her in a portrait. Similarly, switching from seeing a surface merely as a surface to seeing it as a picture, or vice versa, also comes with a noticeable experiential difference. This may happen, for instance, when we move very close to the surface of a painting and look just at a small section of the canvas: we may stop seeing this portion of the surface as depicting anything.

Fifth, the recognitional fold involves the awareness of depth. To get clearer about this observation of Wollheim’s, it is helpful to consider the case of non-figurative paintings, such as Malevich’s *Black Circle* [FIGURE 1]. For cases like this illustrate well how Wollheim conceives of the central element of the recognitional fold and, hence, of our experience of a flat surface as a picture (or ‘representation’):

To see something as a representation is intrinsically bound up with, and even in its highest reaches is merely an elaboration or extension of, the way in

---

<sup>5</sup> At least Wollheim’s earlier view seems to be even stronger, in that he appears to claim that twofoldness concerns the distribution of visual attention, and not only of mere visual awareness (Wollheim 1980 p. 216).

which, when the black paint is applied to white canvas, we can see the black on the white, or behind the white, or level with it. (Wollheim 1965/1974 § 25)

[W]hen seeing-in occurs, two things happen: I am visually aware of the surface that I look at, and I discern something standing out in front of, or (in certain cases) receding behind, something else. (Wollheim 1987 p. 46)

These passages make clear that Wollheim takes not only figurative paintings, but also non-figurative ones to be pictures, which normally give rise to seeing-in (see also *ibid.* p. 62). Indeed, he suggests that figurative seeing-in is just a less basic and more complex form (i.e. an ‘elaboration or extension’) of non-figurative seeing-in.

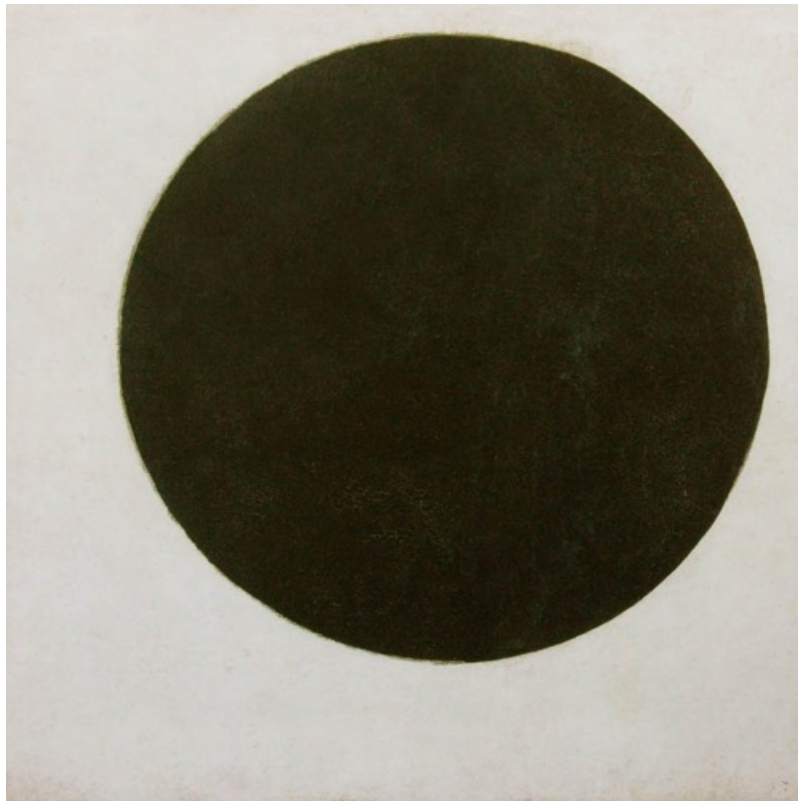


FIGURE 1. Kasimir Malevich, *Black Circle*, 1915, State Russian Museum, St. Petersburg.

Moreover, and more important for our current purposes, the two passages reveal that Wollheim takes the recognitional fold to be, at its core, an *awareness of depth*, that is, an awareness of that dimension of space relative to which things are in front of — or, indeed, level with — other things. He seems perfectly right to claim that the recognitional fold of seeing-in always involves a visual awareness of a three-dimensional arrangement of objects, irrespective of whether these objects are two-dimensional figures (e.g. circles and planes), or whether they are themselves

three-dimensional objects (e.g. trees or animals); and also irrespective of whether the objects are experienced as actually being at different distances relative to the point of view inherent to the picture, or instead as being at the same distance, but with the potential of being at different distances.

This awareness of depth is precisely what distinguishes experiencing a flat surface as a picture from experiencing it merely as a flat surface. If, as part of the recognitional fold, we were not to experience one thing as being in front of another (or as being level with it, but in such a way that it could also be in front of it), then we would not enjoy an experience of something depicted over and above our awareness of the flat surface in question. In particular, Hyman's claim to the contrary is unconvincing (Hyman 2006 pp. 135f. ; Hyman 2003 pp. 681f.). When we look at stick-figure drawings or the silhouettes of people (two of his favourite counterexamples), we usually do not see merely two-dimensional bodies in the surfaces. And even when we do experience the various body parts as being flat and on the same plane, they are at least visually given to us as being in a space which would allow them to be at different distances (e.g. if they were to be rotated, or presented from a different point of view). Similarly, even if depictions of the starry sky stay neutral on whether any of the stars are closer or further away than the others, we are still aware of the depicted sky as allowing for such differences in distance. Seeing-in thus always involves some awareness of depth; and the corresponding pictures are always depictions of depth.

7. The idea that pictorial experience consists in a unified experience of twofold seeing-in, with the five features just listed, is today widely accepted.<sup>6</sup> It also helps us to answer the question about the nature of the normal responses which determine whether something is a picture and what it depicts. Accordingly, something possesses the status of being a picture (rather than that of a mere surface) only if, and partly because, we normally see something in it when looking at it (i.e. normally have a unified experience of it that involves both a configurational and a recognitional fold). Similarly, a picture is a depiction of  $X$  (rather than  $Y$ ) only if, and partly because, what we normally see in it is  $X$  (rather than  $Y$ ).

These considerations about the nature and role of seeing-in raise two important questions, each of which is concerned with one of the two highlighted properties of pictures, their status and their content. The first question is simply about our access to the specific pictorial content of pictures: how are we aware of the fact that a given picture depicts  $X$  (rather than  $Y$ )? The second question, by contrast, ad-

---

<sup>6</sup> In particular, this idea is accepted not only by the proponents of the Experienced Resemblance View and the Imagination View (both to be discussed later on), but also — at least to some extent — by philosophers who are more critical of the orthodoxy, such as Dom Lopes who argues that seeing-in is only one form of pictorial experience among others in order to be able to capture also our standard experiences of illusionistic, naturalistic and actualist pictures (Lopes 2005 ch. 1), or John Kulvicki who likens pictorial experience to perceptual experience as conceived by the sense-data theorists (Kulvicki 2006 chs. 8 and 10).

dresses our access to the status of pictures as depictions. Given that the awareness of depth is central to our recognition of a flat surface as a picture, the question deals more specifically with this kind of awareness: how are we aware of the depth of the depicted scene? For ease of reference, I label these two questions the *Question of Content* and the *Question of Depth*, respectively. Both questions are primarily concerned with the recognitional fold, emphasising the relative importance of this fold for seeing-in. As we will see, the most prominent accounts of seeing-in are unable to answer these two questions in a fully satisfactory manner, generating the need for an alternative elucidation of seeing-in.

8. There has been much debate about the precise nature of twofold seeing-in and, indeed, about how much we can actually say about it. Wollheim is famous for insisting that the configurational and the recognitional fold could not occur on their own, and that we cannot describe their phenomenology by reference to other, independent experiences, notably visual perceptions of surfaces that are not experienced as pictures, and visual perceptions of objects and scenes that are experienced face-to-face rather than by means of depictions (Wollheim 1987 pp. 46f.). As a result, Wollheim concludes that our capacity for seeing-in is a ‘fundamental perceptual [capacity]’ (ibid. pp. 45f.) which is likely to be innate (Wollheim 1986 p. 46) and cannot be much further elucidated. Others have been critical of Wollheim’s conclusion and his arguments in favour of it, and consequently also more optimistic than Wollheim about our prospects of providing a substantial and illuminating account of pictorial experience in terms of seeing-in (Budd 1992/2008; Hopkins 1998 § 1.4; Walton 1992/2008). I turn to two such accounts — the Experienced Resemblance View and the Imagination View — in a minute.

There are two ways in which the configurational and the recognitional fold may be related to each other, assuming that twofold seeing-in constitutes a unified experience. Either the two folds may constitute two different instances of awareness; or they may together form a single instance of awareness. One way of spelling out this distinction is in terms of content.

According to the first option, the unified experiences of seeing-in may involve two contents, one of which presents the surface of the picture as being a certain way (e.g. ‘there are many white areas separated by many thin black lines’), and another which presents the depicted object and its features (e.g. ‘there is a figure reclining on a bed’). Hopkins (2010a p. 20) calls such a view ‘divisive’ since it takes the two contents to be fairly independent of each other.<sup>7</sup>

The second option is to maintain that the configurational and the recognitional fold are nothing but abstractions from a single, complex content which presents the

---

<sup>7</sup> Hopkins uses this term to also describe views — like the theory defended in Gombrich (1960 especially pp. 3ff.) — that deny that pictorial experience consists in twofold seeing-in, and which instead construe pictorial experience as the conjunction or simultaneous occurrence of two distinct experiences.



surface of the picture and the depicted object as standing in a certain relation to each other. Paradigms of this ‘unitary’ view, as Hopkins (*ibid.*) labels it, are the two views just mentioned. While the Experienced Resemblance View argues that a pictorial experience perceptually presents the surface of a picture as looking like the depicted object (e.g. ‘the lines on the surface look like the contour of a reclining figure’), the Imagination View insists that a pictorial experience imaginatively presents the surface as being the depicted object (e.g. ‘the black lines are the contour of a reclining figure’) — or, indeed, that a pictorial experience presents our perceptual experience of the surface as being our perceptual experience of the depicted object (e.g. ‘my seeing the black lines is my seeing the contour of a reclining figure’).<sup>8</sup>

Wollheim’s view on the matter is divisive, in the sense just described. That is, he takes the configurational and the recognitional fold to constitute different instances of awareness or content. This gives rise to the question of how the two folds are supposed to differ in awareness and, in particular, what kind of awareness is assumed to be involved in our visual recognition of what is depicted.

9. The two passages from Wollheim’s writings quoted earlier indicate that he believes that the configurational fold consists in the visual awareness of a *flat* surface, while the recognitional fold involves the visual awareness of *depth*. In addition, he explicitly speaks of ‘seeing’ the black circle in front of the white square and takes seeing-in, as a whole, to be a form or instance of perception (Wollheim 1987 pp. 45f.). This suggests that he assumes that not only the configurational, but also the recognitional fold are perceptual and thus do not differ in the kind of visual awareness that they involve. Taken together, these two elements of Wollheim’s view imply that he conceives of the difference between the two folds of seeing-in solely in terms of depth: they differ because only our visual awareness of what is depicted involves the visual awareness of a third dimension.<sup>9</sup>

But this characterisation of the difference cannot be right. First of all, the surfaces of pictures need not — and often are not — completely flat. Indeed, with many paintings, the relief-like texture of brushstrokes, say, is crucial for depiction (e.g. when van Gogh, in his painting *Wheat Field With Cypresses* [FIGURE 2], uses them to portray the leaves and branches of cypresses in motion). In addition, our perception of the surfaces of pictures locates them in our three-dimensional environment (e.g. in a certain distance from ourselves). Hence, the configurational fold often, if not always, involves awareness of depth as well.

In addition, the proposed characterisation misidentifies the true location of the difference between the two folds. For the configurational and the recognitional fold

---

<sup>8</sup> See sections §§ 13 and 17 for further discussion and references.

<sup>9</sup> Budd (1992/2008 § 5), too, seems to read Wollheim in this way, given that he does not consider the option of accounting for the difference between the two folds in terms of a difference in kind of experience (i.e. perceptual vs. non-perceptual), rather than in terms of a difference in content (i.e. two- vs. three-dimensional, or one kind of depth vs. another).

do not (or not merely) differ in whether they make us visually aware of a third dimension, but instead (or also) in whether they make us *perceptually* aware of the spatial dimensions concerned. While the configurational fold of seeing-in is clearly perceptual, the recognitional fold is not. When looking at Malevich's painting, we do not literally see a black circle as being in front of a white square (or a larger white area that extends beyond that part of the non-figurative scene that is depicted). Nor do we feel any inclination to judge that this is really the case.



FIGURE 2. Vincent van Gogh, *Wheat Field With Cypresses*, 1889, black-and-white detail, Metropolitan Museum of Art, New York.

So, perhaps, a better characterisation of the difference between the two folds is that, while we are perceptually aware of the depicting surface, we are non-perceptually aware of what is depicted. This, however, leaves unanswered the question of what kind of visual awareness is actually involved in the recognition of the depicted objects and features, if not perception — an issue to which we have to return to later on.

10. There is also a second problem for Wollheim's view which does not concern his conception of the difference between the two folds, but instead the fact that he adopts a *divisive* view on the relation between the two folds, namely that they constitute two different instances of awareness or content. The challenge for Wollheim is to identify the two kinds of visual awareness that the configurational and the recognitional fold, respectively, involve.

It should be clear that the two kinds are distinct. One of them makes us aware of something that exists in our environment, right before our eyes; while the other does

not. This is, indeed, partly why one of the folds counts as configurational, and the other as recognitional. Wollheim is also right to insist that the configurational and the recognitional fold cannot be independent of each other in the sense of being able to occur on their own. For, otherwise, the two folds would constitute two distinct experiences that just happen to exist at the same time, rather than one unified experience. So, the challenge is to single out two kinds of visual awareness that are distinct from each other, but none the less cannot be instantiated independently of each other.

Moreover, if the two folds cannot occur on their own, then they cannot be of the same kind as instances of visual awareness that do occur on their own, notably experiences of seeing, recalling or visualising. Thus, when we are looking at a painted portrait of a friend, say, our visual awareness of the depicting canvas cannot be of the same kind as a visual perception, recollection or imagination of a normal, non-depicting surface (e.g. a painted wall).<sup>10</sup> Similarly, our visual awareness of the person portrayed cannot be of the same kind as a standard (i.e. ‘face-to-face’) visual perception, recollection or imagination of such a person. The two kinds of visual awareness in need of identification therefore have to be different from seeing, recalling, visualising or any other kind of visual awareness that enjoys distinct existence. This means that the two folds of seeing-in must constitute two distinct, but interdependent *sui generis* kinds of visual awareness, which can occur only in the context of pictorial experience.

The plausibility of this idea is closely linked to the issue of whether something more illuminating can be said about the nature of these two kinds of visual awareness. Their postulation seems already to be rather ad hoc, given that it is motivated solely by the need to account for the twofoldness of seeing-in, and given that the two kinds of visual awareness are assumed to play no role outside of seeing-in. But if, in addition, their nature remains completely unelucidated, their introduction loses not only any explanatory power attributed to it, but also becomes very difficult to assess. If it remains ill-defined which two kinds of visual awareness a divisive view like Wollheim’s actually postulates, we should in fact reject this view for being unclear.

11. Wollheim seems to think that this charge is misguided because it overlooks the fact that twofoldness introduces such a complexity into seeing-in that, as a result, the configurational and the recognitional fold turn out to be ‘incommensurate’ with simpler visual experiences that enjoy independent existence (Wollheim 1987 p. 47). What this means is that, while there may very well be certain experiential similarities

---

<sup>10</sup> Walton (1992/2008 p. 134) seems to maintain that the configurational fold can be of the same kind as a normal perception of a flat surface. The underlying thought appears to be that even normal perception is always penetrated by the thought ‘I experience X’, thus ensuring that it involves the same kind of cognitive penetration as seeing-in. See §§ 17-18 for a discussion of Walton’s view.

between seeing-in and other visual experiences, we are not in a position to inquire into and discover their nature. In particular, it is futile to try to describe the two folds of seeing-in by comparing them to standard visual perceptions (or recollections, imaginations, etc.) of surfaces and objects. So, according to Wollheim, it is the very nature of seeing-in that prevents us from saying more about the two kinds of visual awareness involved in it.

However, Wollheim's claim about the incommensurability of seeing-in does not enjoy much plausibility. First, as Budd (1992/2008 p. 200) notes, that seeing-in is a complex whole does not show that none of its parts (i.e. its folds) can be compared with other visual experiences. Second, assuming that we are generally able to introspect and describe similarities in awareness between different kinds of experience, that we could not report any similarities between seeing-in and other experiences would be a good sign that there are no such similarities. But this conclusion would seem to be too strong, given that seeing-in shares at least some properties of other visual experiences — notably their visual character. Third, it appears that we are actually able to introspectively note and characterise some of the similarities between (the two folds of) seeing-in and the other experiences. We do not seem to have difficulties to recognise that the configurational fold comes very close to a standard perception of a painted surface (e.g. in making us visually aware of a flat expanse with marks of paint on it), or that the recognitional fold is much alike a visual experience face-to-face (e.g. in making us visually aware of some voluminous object or scene).

12. But even if it is possible to identify and describe the various similarities (e.g. in content, attitude, functional role, etc.) between the two kinds of visual awareness involved in seeing-in, on the one hand, and standard kinds of visual experience, on the other, this is unlikely to fully remove the challenge to divisive views in general, and to Wollheim's version in particular.

Knowing that the configurational fold resembles a normal perception of a flat surface in visually presenting a marked surface as really being a certain way and eliciting a corresponding perceptual belief does not tell us anything about the nature of that fold. That it is *like* standard instances of perception might suggest that it itself *is* a standard instance of perception. But this conclusion would be wrong precisely because the configurational fold cannot, unlike normal perception, occur without the recognitional fold. And the recognition of the resemblance does not suggest any better alternative conclusions.

Similarly, noting that the recognitional fold is very similar to visualising in that both visually present some objects without (necessarily) making any claim about their existence or real nature and without (normally) giving rise to matching beliefs, might at best provide some evidence for believing that the fold is an instance of visualising.<sup>11</sup> But this verdict would, again, be wrong because of the essential

---

<sup>11</sup> I do believe that visualising, taken on its own, never makes a claim about how reality is like

dependence of the recognitional fold on the configurational one.

Unless the challenge of providing a proper elucidation of the nature of the two kinds of visual awareness said to be involved in seeing-in is met, we should prefer unitary accounts over the divisive views of Wollheim and others. The two main candidates on the market are, as already mentioned, the Experienced Resemblance View and the Imagination View. However, neither of them can satisfactorily account for both of the two facts discussed at the beginning. While the Imagination View fails to explain our recognition of what a picture depicts, the Experienced Resemblance View cannot explain our awareness of flat surfaces as pictures, notably pictures of voluminous objects or scenes.

13. Let us begin with the *Experienced Resemblance View*. Proponents of this account maintain that seeing-in has to be elucidated in terms of an experience of resemblance. Its main claim is that recognising surfaces as pictures essentially involves experiencing them as resembling some other objects, namely the depicted ones. Accordingly, an object counts as a picture only if it gives rise to such an experience of resemblance; and it depicts those entities which it is experienced as resembling.

The resulting view is *unitary* since it interprets the two folds of seeing-in as being combined in a single instance of awareness, with something like ‘the surface looks like X’ (where ‘X’ describes the depicted objects or scene) as its content. And the view postulates merely an *experienced* resemblance, rather than a real one, since it does not require that the depicting surface and the depicted objects are really similar to each other. Indeed, the depicted objects need not even exist, in which case there could not be any relation of resemblance in the first place.

Different philosophers have presented different accounts of the resemblance relation that we are assumed to experience between depicting surfaces and depicted objects.<sup>12</sup> But their views have in common that they focus on two-dimensional shapes as the focus of similarity. In other words, they argue that pictures and what they depict are experienced as resembling each other with respect to certain two-dimensional shapes.<sup>13</sup>

14. The two relevant sets of two-dimensional shape may be labelled surface shapes and outline shapes.<sup>14</sup> *Surface shapes* pertain to flat surfaces, such as those surfaces

---

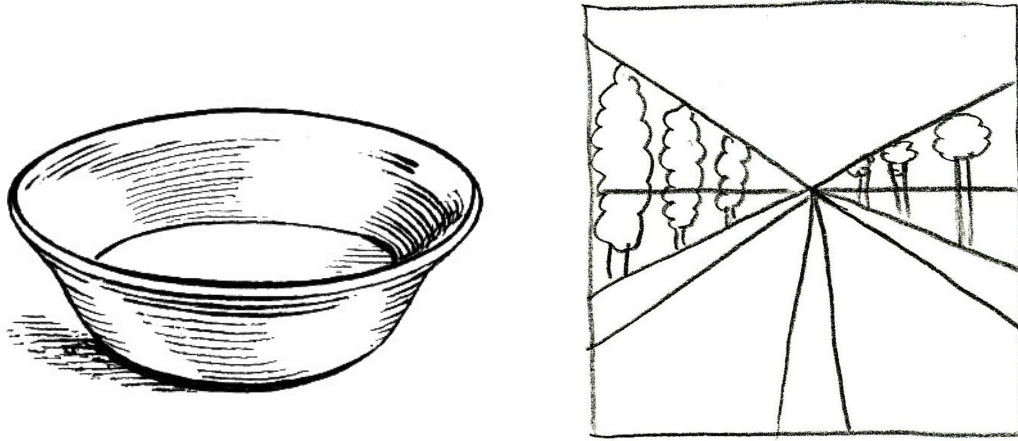
and never leads to belief in a rational person (Dorsch 2015b), but, for the current argument, it suffices if this claim holds in typical circumstances.

<sup>12</sup> See Peacocke 1987, Budd 1991/2008, Hopkins 1998, and Hyman 2006. Note, however, that the last of these philosophers does not account for depiction in terms of experienced resemblance and thus does not endorse the Experienced Resemblance View.

<sup>13</sup> Hyman (ibid.) argues that the two are also experienced as being similar in certain aspects of colour. Here, it suffices to focus on experiences of resemblance in two-dimensional shape because of their close link to depth awareness.

<sup>14</sup> In the light of the earlier observation that painted surfaces are hardly ever entirely flat and

that depict something. They consist in the shapes formed by the lines and coloured patches arranged on the surface in question. For example, the black lines drawn on a white sheet of paper may delineate an elliptical or a triangular area [FIGURES 3 AND 4].



FIGURES 3 AND 4. Unknown Artists, *A Bowl & A Road with a Median Strip.*

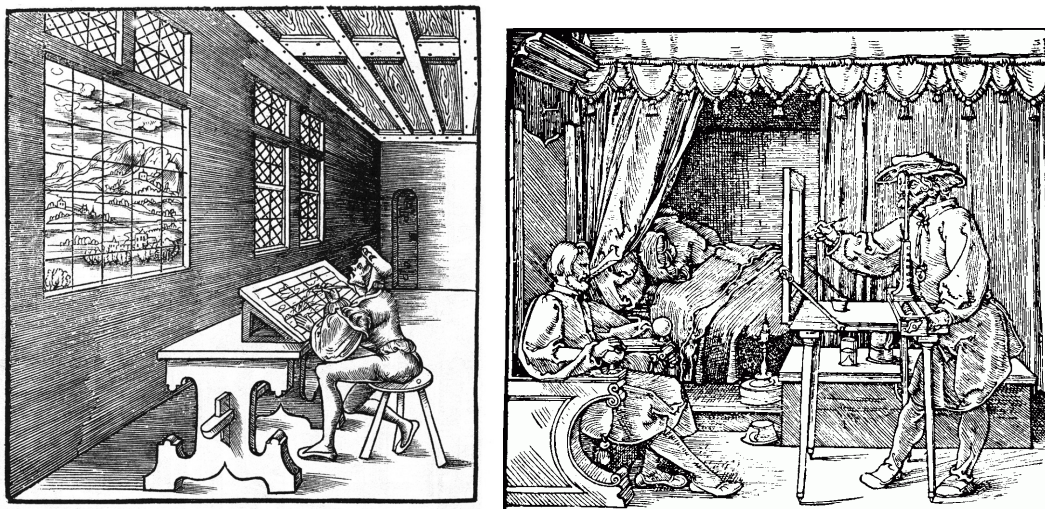
*Outline shapes*, by contrast, pertain to voluminous objects, such as those objects that pictures depict. They are identical with the shapes that one would draw on a transparent plane (e.g., a pane of glass), aligned orthogonally to the line of sight, in order to trace the outlines of such objects as seen through that plane. Paradigm examples of apparatuses for tracing outline shapes are the various perspective frames and machines devised by Leon Batista Alberti, as shown in the two woodcuts by Count Johann and by Albrecht Dürer, respectively [FIGURES 5 AND 6]. When the circular rim of a bowl is seen from the side through such an apparatus, its outline shape turns out to be elliptical; while the outline shape of a road that disappears into the distance will be triangular. We can thus understand two-dimensional outline shapes as abstractions from three-dimensional volumes that discard the dimension of depth.<sup>15</sup> Because of this abstraction from the third dimension, outline shapes (unlike surface shapes) turn out to be relative to a certain point of view (i.e. to a combination of a certain location in space and a certain direction of gaze). For instance, when the rim of the bowl is seen from above, its outline shape will be

---

two-dimensional, the surface shapes are, strictly speaking, also outline shapes, relative to a frontal point of view on the surface. But this complication can be safely ignored in what follows.

<sup>15</sup> See Hopkins 1998 p. 115. With some qualifications, this characterisation of outline shapes captures the core idea of Hopkins' own notion of 'outline shape' (the plane is located between object and viewer), of Hyman's notion of 'occlusion shape' (the plane is located behind the object) and even of Budd's and Peacocke's notion of 'visual field shape' (the plane is identical with what is left over if we abstract the dimension of depth from the three-dimensional scene that we experience; see Budd 1992/2008 p. 208 and Budd 1991/2008 p. 221).

circular, rather than elliptical.<sup>16</sup>



FIGURES 5 AND 6. Count Johann, *Window Grid*, 1531. — Albrecht Dürer, *An Artist Drawing a Seated Man*, 1525.

Now, according to the Experienced Resemblance View, what is central to seeing-in is the complex visual experience of the surface shapes of the depicting surface as looking like the outline shapes of the depicted objects. For example, when we look at a drawing that depicts a round bowl from a certain point of view, we visually experience the relevant elliptical shape on the sheet of paper as resembling the elliptical outline shape that the rim of such a round bowl would possess relative to the point of view in question [FIGURE 3]. Similarly, in the case of a picture of a road disappearing into the distance, the relevant triangular shapes formed by the lines on the picture's surface are visually experienced by us as being similar to the triangular outline shapes that the lanes and the median strip of a straight road possess relative to the perspective in question [FIGURE 4].

15. There are some open questions about the nature of this visual experience of resemblance. The only thing that is clear is that we see the surface shapes, but not the outline shapes. For we see the picture, but not what is depicted, given that only the former, but not the latter is experienced as being there before our eyes. As things appear to us, there is only a picture of a bowl in front of us, but no bowl.

---

<sup>16</sup> In this respect, outline shapes resemble the apparent or variable shapes that figure in our perceptual experiences of the constant shapes of objects. A coin perceptually appears to look round (i.e. to have a certain constant shape) irrespective from which perspective we look at it. But our perceptual awareness of the coin and its constant roundness also involves, in some way or another, the variable property of being elliptical, which changes from one point of view to another. This variable shape property is merely apparent since the coin does not perceptually appear to possess it and since, relatedly, our perceptual experience of the coin neither inclines, nor entitles us to ascribe it in belief to the coin.

Seeing-in is thus said to be more like seeing a person and recognising her resemblance to another one who is far away or does not even exist, rather than like noticing a resemblance between two people while seeing both at the same time. This raises the question of how we are aware of the outline shape of the bowl or of whatever is depicted, if not perceptually.

It is surprising how difficult it seems to provide a satisfactory answer to this question. Hopkins, for instance, maintains that our awareness of the depicted consists in a ‘thought’ about its visual appearance, ‘in the light of which’ we see the picture’s surface, but he also acknowledges that ‘there are limits to what I think can illuminatingly be said about experienced resemblance itself (Hopkins 2003 p. 156). Apart from the problem that it is unclear whether ‘thinking’ may indeed be understood in such a wide sense as to include visually experiencing, we are not really any closer to an answer to our question until it is specified what kind of ‘thought’ is at issue.

Perhaps, when the picture of a bowl depicts our own bowl at home and we recognise this fact, our awareness of the bowl’s outline shape might be a product of memory. But looking at the picture does not include episodic memory, that is, the recollection of one particular of our past perceptions of our own bowl. We do not have the appearance of our bowl before our mind, as we saw it in the past. Hence, the form of memory involved could at best be generic, that is, the kind of non-occurrent visual memory that comes with, or consists in, the capacity to reidentify our particular bowl on the basis of visual perception. Such generic memory might indeed give rise to, inform or become manifest in our visual awareness of the bowl’s outline shape as part of seeing-in. But unless we know more about the nature of this kind of awareness, the Experienced Resemblance View is still incomplete and difficult to assess.

In addition, pictures need not be of any objects that existed in the past or exist in the present. Hence, the picture in our example may just depict a bowl of a certain type (i.e. one with certain qualities), without depicting any particular bowl (i.e. one with a specific numerical identity). In this case, the idea of recalling the bowl and its outline shape relative to a certain point of view would not be an option. As several proponents of the Experienced Resemblance View have noted, we also cannot treat our awareness of the bowl as an instance of the visual imagination. Pictorial experience is not a matter of seeing a picture’s surface and simultaneously visualising the depicted, partly because the resulting experience would lack the unity characteristic of seeing-in and, in particular, would not include experiencing the depicted by means of, or on the basis of, seeing the picture.<sup>17</sup>

Moreover, even if we get a sufficient grasp of the awareness of outline shape supposed to be involved in seeing-in, it is still undecided whether our awareness of the relation of resemblance between the two sets of shapes is perceptual or not.

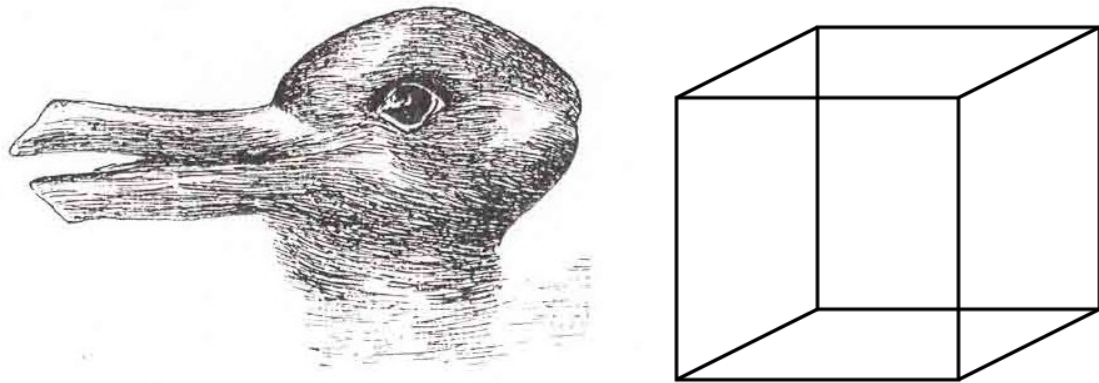
---

<sup>17</sup> See Budd 1992 and Hopkins 1998 pp. 20ff. for this and other arguments against the Imagination View, which I discuss in more detail in § 17.



That we see the two people whom we experience as resembling does not necessarily mean that we see them as being similar to each other. We might also just imagine there being a resemblance between the two on the basis of seeing them. Again, proponents of the Experienced Resemblance View are not very explicit on this issue. True, as observed earlier, they do not assume that the depicting surfaces and the depicted objects resemble each other. But claiming that this resemblance is merely experienced leaves it open whether it is experienced perceptually or in a different manner. Perhaps it is best to reject the perceptual option, simply because this could mean that seeing-in turns out to be systematically illusory, given that there may very well be no real similarity.

16. In any case, the most important issue for our current purposes is whether the Experienced Resemblance View can provide satisfactory answers to the two questions introduced in § 7. With respect to the Question of Content, this seems to be the case (with the proviso that the issues raised in the last section can be resolved). We recognise that a given picture depicts  $X$  (rather than  $Y$ ) by experiencing its surface shapes as resembling the outline shapes of  $X$  (rather than those of  $Y$ ). In other words, which pictorial content we experience a picture as having is constrained by what we experience to be similar to the picture's surface with respect to two-dimensional shape.



FIGURES 7 AND 8. Unknown Artists, *Duck-Rabbit Figure & Necker Cube*.

Of course, there are ambiguous pictures — such as the duck-rabbit drawing included in Wittgenstein (1953/2009 part II § xi), or the Necker cube line drawing [FIGURES 7 AND 8] — whose surface shapes can be experienced as looking like the outline shapes of several, often incompatible things. But even then, there is a strict limitation on what we can recognise the pictures as depicting; and this limitation is explained by our capacity to recognise similarities between surface and outline shapes. For instance, we cannot see the duck-rabbit picture as a depiction of a horse, say, because the shape of the outline on the drawing's surface does not look

like the outline shape of a horse. More generally, among the kinds of object that we are capable of visually identifying, hardly any share exactly the same outline shape relative to a given point of view.<sup>18</sup>

The Experienced Resemblance View fares, however, much worse with respect to the Question of Depth (which, it must be added, it was not explicitly designed to answer). Since, as noted earlier, pictures are always pictures of three-dimensional arrangements of objects, recognising a flat surface as a picture requires an awareness of depth with respect to what is depicted. But the proposed experience of resemblance is limited to the two-dimensional outline shapes of the depicted objects. Hence, it cannot constitute or contain the awareness of depth that is involved in the recognitional fold of seeing-in. As a result, the Experienced Resemblance View — at least in its current form — does not have the resources to capture this central aspect of the recognitional fold and is thus unable to fully capture the nature of pictorial experience.

17. The Imagination View does much better on this count. At the heart of this view is the idea that our awareness of what is depicted is imaginative. After all, the recognitional fold is like imagination, and unlike perception, in that we do not experience the depicted objects as really being there in front of our eyes.

The kind of imagining in question, however, can be neither visualising, nor propositional imagining (i.e. imaginative thinking). The main reason for this is again that, as argued earlier, a unitary account of seeing-in is to be preferred over a divisive one. But while seeing-in should thus be taken to involve two interdependent folds, instances of visualising and propositional imagining can occur independently of other mental episodes, notably perceptions of flat surfaces. In addition, already the fact that seeing-in constitutes a single, unified experience speaks against the identification of the imaginative element with visualising or propositional imagining, given that instances of the latter do not form such unities together with simultaneous perceptions of suitable flat surfaces.<sup>19</sup> Besides, propositional thought — even if it is about visual appearances — does not show the visual and spatially perspectival character of our awareness of what is depicted.

Instead, Kendall Walton, the most prominent defender of the Imagination View, has proposed to understand seeing-in as the experience of imagining one's perception of the picture's surface to be a perception of what is depicted. Accordingly, when we have a pictorial experience of a portrait of a friend, say, we are said to imagine our actual perception of the portrait to be a perception of our friend. As part of imagining this identity of two perceptions, we also imagine the identity of the respective objects of perception. That is, in our example, we imagine the portrait

---

<sup>18</sup> See also the discussion in Hopkins 1998 § 5.7 which argues that, even if many different things share outline shapes, this does not undermine the success of the Experienced Resemblance View in explaining what the difference is between seeing *X* and seeing *Y* in a given picture.

<sup>19</sup> See Dorsch 2015a, as well as Budd 1992 and Hopkins (1998 pp. 20ff.) for discussion.

to be our friend. This becomes evident, for instance, in the fact that it would be perfectly natural for us to say, while pointing to the portrait, ‘there is our friend’ or ‘here we see our friend’ (Walton 1990 pp. 215ff.).

This imagination likewise extends to the features of the objects concerned. For instance, when we are scanning the relevant marks of paint on the canvas with our eyes, we imagine this very experience to be a scanning of the face of our friend. And we imagine certain features of the canvas that we thereby see (e.g. some colours or shapes) to be (closely related to) features of that face. Walton argues that, in this way, the visuality of our awareness of the depicted can be preserved, given that the ‘vividness’ — for example, the continuity and detailedness — of our real perception of the canvas transfers to the imagined perception of the friend (ibid. ch. 8, especially p. 296 and § 8.2).

Moreover, imagining this identity between the real perception of the picture’s surface and the imagined perception of the depicted is meant to ensure that the two folds of seeing-in are united in a single experience with a single (though complex) content, with respect to which the configurational and the recognitional fold are merely abstractions.<sup>20</sup> More specifically, Walton’s idea is that, in imagining our perception of the portrait to be a perception of our friend, the imaginative thought of our friend penetrates our perception of the portrait and changes its character and content. In other words, seeing-in involves, for Walton, the cognitive penetration of the visual perception of the picture’s surface by an imaginative thought about what is depicted.<sup>21</sup>

18. One problem with the resulting Imagination View is, however, that it is difficult to motivate the classification of our awareness of the depicted as an imaginative thought. First of all, there are good reasons to believe that imagining is generally susceptible to the will (McGinn 2004 pp. 12ff. and Dorsch 2012 chs. 13f.). But, apart from ambiguous pictures, we cannot actively bring about changes in what we take

---

<sup>20</sup> See Walton 1990 pp. 295 and 301. In fact, Walton acknowledges that the configurational fold can occur on its own (see footnote 10). But it is still true for him that the recognitional fold requires the configurational fold, given that the imagining involved in seeing-in is *de re* with respect to the perception of the picture (Dorsch 2015a). Moreover, as already mentioned in § 6, the experience of a surface as a mere surface and the experience of the same surface as a picture are very unlikely completely indistinguishable from the inside.

<sup>21</sup> See Walton 1990 pp. 295 and 300. Indeed, Walton thinks that visual experience — including perception — is in general cognitively penetrated by the self-referential thought ‘I am experiencing X’ (ibid. p. 295 n. 3). And he likens his own view of seeing-in to Wittgenstein’s notions of ‘seeing as’ and ‘aspect seeing’ as fusions of seeing and thinking (ibid. p. 295 n. 2, and p. 351), which suggests that his view might be closer than initially recognisable to the one that I will put forward in §§ 22ff. In any case, Walton’s conception of seeing-in as consisting in cognitively penetrated perception supports his repeated insistence that seeing-in is an instance of perception, albeit of imaginative perception (ibid. p. 302 and Walton 2002 n. 37). As I discuss in Dorsch (2015a), the account of pictorial experience defended in O’Shaughnessy (2003 pp. 346ff.) is very similar to Walton’s (see also O’Shaughnessy 2012).

a given picture to depict, that is, what we imagine ourselves as seeing when looking at the picture. Hence, our awareness of the depicted cannot really be imaginative. In addition, it is questionable whether it can be an instance of thought, for the reason already mentioned: namely that our awareness of the depicted is visual and spatially perspectival, while thought is not.

But even if we accept that seeing-in combines perception and imaginative thought, this combination could not involve cognitive penetration. Cognitive penetration changes the perceptual experience concerned with respect to how we perceive the object in question as being (Stokes 2013 § 1). When it suddenly strikes us that the object that we are currently seeing is an engine, say, this thought is likely to enter into our perception of the object. As a result, we begin to experience it as being an engine, which we did not before. But we do not, in any way, *experience* the canvas before us as being our friend when having a pictorial experience of the portrait concerned. We might perhaps entertain the thought that the canvas is our friend. But this does not influence our perception of the canvas in such a way that we visually experience it as having the property of being our friend. In fact, how we perceive the canvas as being (e.g. which colours, shapes and textures it possesses) has not changed. Part of the underlying explanation of this difference is that, while cognitive penetration involves only one object of awareness, seeing-in involves two distinct ones (i.e. the picture and the depicted). For there is no reason for us to expect that thinking of *one* object as being a certain way has any impact on how we see *another* object as being.<sup>22</sup>

19. With respect to our two questions, the Imagination View has, finally, the complementary problem to the difficulty faced by the Experienced Resemblance View. The Imagination View promises a rudimentary answer to the Question of Depth — at least as long as we ignore the problems that have been already noted in relation to the idea of understanding seeing-in as a combination of seeing and imagining.

The view claims that we are aware of something as a picture by imagining our experience of it to be an experience of the depicted, which includes imagining the picture's surface to be the depicted. As part of this imaginative identification, we imagine the depicted arrangement of objects to be three-dimensional in one way or another — say, as consisting of a black circle in front of a white plane, or as involving some voluminous objects, such as our friend and his surroundings. In other words, we are aware of the depth of the depicted scene by imagining it to be three-dimensional; and we imagine it to be this way as part of imagining (our

---

<sup>22</sup> There is also the problem that cognitive penetration seems to require sameness in attitude or commitment. Accordingly, while judgemental thoughts are limited to entering into perceptual experiences, imaginative thoughts are restricted to merging with imaginative experiences. For instance, when we notice the visual resemblance between a bouncer at a club and a gorilla and, as result, see him under the concept of a gorilla, we do not see him *as being a gorilla* (because we do not believe him to be a gorilla), but rather see him *as looking as a gorilla* (because we do take him to be visually similar to a gorilla). See Dorsch (2015a) for further discussion.

perception of) the two-dimensional surface of the picture to be (a perception of) a three-dimensional arrangement of objects.

By contrast, the Imagination View has difficulties to answer the Question of Content. The challenge is to explain why we imagine perceiving *X*, rather than *Y*. For example, when we look at the portrait of our friend and scan the canvas with our eyes, why do we imagine this to be a perceptual experience of our friend's face, and not, say, of some animal or a tree? Walton's answer to this question is twofold.

First, he proposes that there are certain principles or rules (called 'principles of generation' by Walton) which determine what a given picture depicts (Walton 1990 p. 110). These principles are very similar to those that fix what words or sentences in a certain language mean, or what props in games of make-believe stand for. In particular, they are, at least to a considerable extent, intersubjective (*ibid.* pp. 23, 68 and 301f.), as well as natural and non-arbitrary, rather than conventional and stipulated (*ibid.* pp. 23, 301 and 351f.). Besides, given that they determine what a picture depicts, they also dictate what it is appropriate for us to imagine when looking at the picture and recognising it as a depiction (*ibid.* pp. 215ff. and 293ff.).

Second, we normally imagine the depicted in accordance with these principles because we are prompted to do so by perceiving certain properties of the picture's surface (Walton 2002 p. 32). Typically, this transition from the perceptual awareness of the picture to the imaginative awareness of the depicted is involuntary and non-inferential (Walton 1990 pp. 216 and 311). Moreover, it need not — and usually does not — involve explicit knowledge of, guidance by or reflection on the principles in question (*ibid.* pp. 139, 185f. and 216). Instead, we implicitly understand and accept the principles simply by being disposed to imagine whatever is depicted in response to seeing the relevant properties of the picture's surface (*ibid.* p. 216). As Walton writes:

Often it just strikes us that, given the words of a novel or the paint on a stretch of canvas, such and such is fictional [i.e. to be imagined]. (*ibid.* p. 139)

Even assuming that our spontaneous understanding of what is to be imagined can be explained in terms of rule-following, this reply to the Question of Content is at best a partial answer. For there still remains a set of crucial questions: which properties of the surfaces of pictures determine, and prompt us to imagine, what they depict; and what kinds of link, expressed by the principles, obtain between these properties and what is depicted?

20. Walton leaves these questions open and, furthermore, stays neutral on whether our implicit knowledge of the principles is innate or acquired, and on whether the principles are open to changes (*ibid.* pp. 301f.). But he seems to be sympathetic to the natural idea that what is at work here is the awareness of visual similarities between the picture and the depicted, notably in two-dimensional shape (Walton 2002 p. 31). In fact, the suggestion that the principles are concerned with similarities

in surface and outline shape fits also very well Walton's observation that, as part of seeing-in, we often scan the outlines on the canvas and imagine of this experience that it is an experience of scanning the outlines of whatever is depicted.<sup>23</sup>

But if it were indeed the case that our imaginative awareness of the depicted is informed by a prior awareness of the resemblance between the picture and the depicted, then it seems that all the elements needed for seeing-in are already there before we start to imagine anything. For, although the awareness of resemblance might not be all that there is to seeing-is, it is at least sufficient for the presence of the configurational and the recognitional fold. In particular, being aware of how what is depicted resembles the marks on the picture's surface in the relevant respects is already sufficient for being aware of the depicted and its features; and there is no need for any further kind of awareness, imaginative or otherwise. In short, imagining the depicted seems already to presuppose seeing-in.

21. Walton might reply that the awareness of resemblance in question is merely subpersonal, while the recognitional fold occurs on the personal level. However, it should be clear that it is often, if not always, possible that we come to personally experience the relevant similarities. In particular, people enjoying a pictorial experience are generally able to report, or point to, the similarities in question when prompted to do so, which again requires that they become aware of those resemblances on the personal level. But if they do start to experience the similarities between the picture and the depicted, there is once more no need to some subsequent imaginative awareness of the latter. Hence, the Imagination View could at best insist that seeing-in involves imagining only with respect to those cases in which our awareness of resemblance actually remains subpersonal. But, apart from the fact that such cases would be rather rare (if they did occur at all), this would lead to the implausible consequence that there are more than one kind of seeing-in, namely one involving imagining, and another involving the experience of similarity in two-dimensional shape.

Unless the Imagination View comes up with a better explanation of how what we imagine while having pictorial experiences is directly constrained by how the surfaces of the pictures concerned are like, it is unable to answer the Question of Content without giving decisive ground to the Experienced Resemblance View. Hence, it is best, at least for the time being, to conclude that whether we imagine seeing *X* or *Y* when looking at a picture is normally determined by whether we see *X* or *Y* in the picture (which again may — but need not — depend on whether we experience the picture as resembling *X* or *Y*); and not the other way round. As Walton rightly notes, 'it seems undeniable' that someone looking at the picture of a ship, say, typically imagines him- or herself looking at a ship (ibid. p. 217). But this observation is compatible with the view that this form of imagining is not part

---

<sup>23</sup> See Walton 1990 § 8.2 for a detailed description of how our imaginative engagement with the depicted may be grounded in, and correspond to, our perceptual engagement with the picture.

of, but instead based on and informed by seeing-in.

22. Both the Experienced Resemblance View and the Imagination View fail to satisfactorily answer one of the pair of questions introduced in § 7. But their shortcomings and, more important, also their strengths are complimentary. In particular, while the Experienced Resemblance View has a satisfactory reply to the Question of Content, the Imagination View at least suggests an answer to the Question of Depth. So, perhaps it is possible to combine these two elements and maintain that seeing-in involves both an experience of resemblance in two-dimensional shape and a non-perceptual awareness of depth.

This latter awareness should thereby not be understood as imaginative, for the reasons spelled out earlier — most notably because imagining is subject to the will and cannot cognitively penetrate a perceptual experience of a flat surface in such a way as to give rise to an experience of seeing-in. Accordingly, seeing something three-dimensional in a two-dimensional surface does not amount to imagining the third dimension of the objects, the outline shapes of which we experience to be similar to the respective surface shapes.



FIGURE 9. Ronald C. James, *Dalmatian Dog*, 1965.

Instead, I suggest that we conceive of seeing-in as being an instance of *seeing an aspect*.<sup>24</sup> The aspect in question is the property of having the visual appearance of a

---

<sup>24</sup> Voltolini (2012, 2015) also defends the idea that seeing-in involves seeing an aspect, although

three-dimensional arrangement of two- or three-dimensional objects (i.e. the visual appearance of the depicted scene). As part of seeing this aspect, we experience the outline shapes of these objects as looking like certain surface shapes on the canvas. But seeing this aspect also involves being visually aware of depth. More specifically, we see the aspect by seeing the marks on the picture's surface — or, perhaps more neutrally, the points on, or parts of, that surface — as being visually organised in a certain three-dimensional manner. In the well-known picture of a Dalmatian dog [FIGURE 9], for instance, we perceptually experience the two-dimensional black areas on the white sheet of paper as being arranged the three-dimensional shape of a dog in an autumn landscape.



FIGURES 10 AND 11. Rembrandt Harmenszoon van Rijn,  
*Posthumous Portrait of the Preacher Jan Cornelisz. Sylvius*, ca. 1645,  
 National Gallery of Art, Washington & *Sketch for the Posthumous Etching  
 of Jan Cornelisz. Sylvius*, ca. 1646, British Museum, London.

How we see the marks as being visually organised depends in part on the relevant surface shapes. That is, we see the marks as forming a three-dimensional arrangement of objects, the outline shapes of which are experienced by us as resembling the surface shapes. In the example, some of the surface shapes suggested by the black areas look like the outline shapes of the dog. Other properties of the marks on the picture's surface may, too, contribute to the visual organisation of the aspect

---

he maintains that aspect perception already occurs on the level of the configurational fold. A comparison of his to my view has to wait for another occasion.



as well, notably colours and textures. Indeed, examples of inflection can perhaps be accounted for by reference to the fact that certain material qualities of the surface may contribute to the three-dimensional visual appearance that constitutes the aspect noticed in seeing-in — for example, when we see the hands of Jan Cornelisz Sylvius, depicted in his two portraits by Rembrandt, as being made of paint or ink, rather than flesh [FIGURES 10 AND 11].<sup>25</sup> However, experienced resemblance and experienced inflection are likely to be only two elements among many that are involved in seeing the aspect under consideration.

23. The resulting account — the Aspect View of perceptual experience — gets support from the fact that our awareness of what is depicted shows many of the features characteristic of experiences of aspects. Indeed, the idea that seeing-in involves seeing an aspect is not new:<sup>26</sup>

When we are looking at an object we sometimes see that it has not changed while we have been looking at it and yet the way in which we see it has changed: we see it differently, although we see that it is no different from how it was. For example, we might pass from seeing a puzzle-picture as mere lines to seeing it as containing a depiction of a face, [...] from seeing the ambiguous duck-rabbit figure as a duck-picture to seeing it as a rabbit-picture, from seeing [a certain drawing] as an arrow (a sign in the form of an arrow, pointing in a certain direction) to seeing it as depicting a bird's foot [...]. In each case we can be said to notice an *aspect* of what we are looking at. (Budd 1987 p. 1)

In this passage, Budd represents Wittgenstein's view on the matter, which is likely also his own. According to this view, both the switch from seeing something as a mere surface to seeing something as a depiction (or vice versa) and the switch from seeing something as a depiction of  $X$  to seeing it as a depiction of  $Y$  are changes in whether, or which, aspects are seen by us. This might seem at odds with my suggestion that the aspect involved in seeing-in is the property of having the visual appearance of a three-dimensional scene, rather than the property of being a depiction.

But there is no real tension here. It is part of the twofoldness of seeing-in that we recognise a flat surface as a picture by becoming visually aware of a three-dimensional scene over and above, and on the basis of, perceiving the two-dimensional surface. The perception of the surface (i.e. the configurational fold) is concerned with the visible features of the marks on the surface, such as their locations, shapes, colours, textures, and so on.<sup>27</sup> As a result, this visual awareness

---

<sup>25</sup> See, for instance, Hopkins (2010a) for a general discussion of inflection.

<sup>26</sup> The omitted parts of the passage make clear that Wittgenstein (and Budd) also think(s) that experiencing a similarity amounts to seeing an aspect. This strengthens my suggestion that the experience of resemblance in two-dimensional shape is part of seeing the aspect of a three-dimensional scene.

<sup>27</sup> It may also involve the perception of more higher-level properties, most notably grouping

stays constant during the various switches just described.<sup>28</sup> That is, it does not make us aware of the aspect which we notice when coming to recognise something as a picture. Instead, the awareness of this aspect pertains to the visual awareness of the depicted (i.e. the recognitional fold). And visually experiencing what is depicted just consists in visually experiencing the appearance of a three-dimensional arrangement of objects.

The conclusion should therefore be that recognising something as a picture involves seeing an aspect only in so far as, and because, it involves visually experiencing a three-dimensional scene (over and above seeing a flat surface). What changes during the switches in seeing-in is whether we are aware of such a scene, or which one we are aware of. The aspect that we notice by means of seeing-in is thus the property of having the visual appearance of a three-dimensional scene.

24. These considerations show that seeing-in satisfies the basic characterisation of seeing an aspect, which the passage quoted from Budd's text starts off with. The experience of seeing an aspect is defined by the fact that, while switching from seeing one aspect to another, from seeing no aspect to seeing one, or from seeing an aspect to seeing none, we do not see the object in question as changing, although our experience of it changes. This is exactly how it is with pictures and seeing-in. We do not experience the picture's surface as changing when shifting from seeing it merely as a flat surface to seeing it as a picture. Nor do we see any changes in the lines of the duck-rabbit drawing while moving back and forth between seeing a duck in it and seeing a rabbit in it.

This is due to the fact that our perception of the locations, shapes, colours and textures of the marks on the surface stays constant throughout, irrespective of whether we are also aware of the visual appearance of a three-dimensional scene. In other words, the change that we experience during the switches pertains to the recognitional fold, not the configurational one. Indeed, it is a change that occurs in our mind, not in the object — which is in line with the fact that being a picture and having the visual appearance of a three-dimensional arrangement of objects are response-dependent properties.

In addition, seeing-in possesses other important features that are central to aspect perception. To start with, seeing-in is an experience which is both visual and immediate and, as such, does not involve any conscious form of reasoning or inference. The same is true of other instances of aspect perception, such as the experience of grouping phenomena. Then, seeing-in allows for the possibility of error. We may fail to recognise something as a picture; we may hallucinate a depicting surface; and our experience may mislead us about what a given picture depicts (e.g.

---

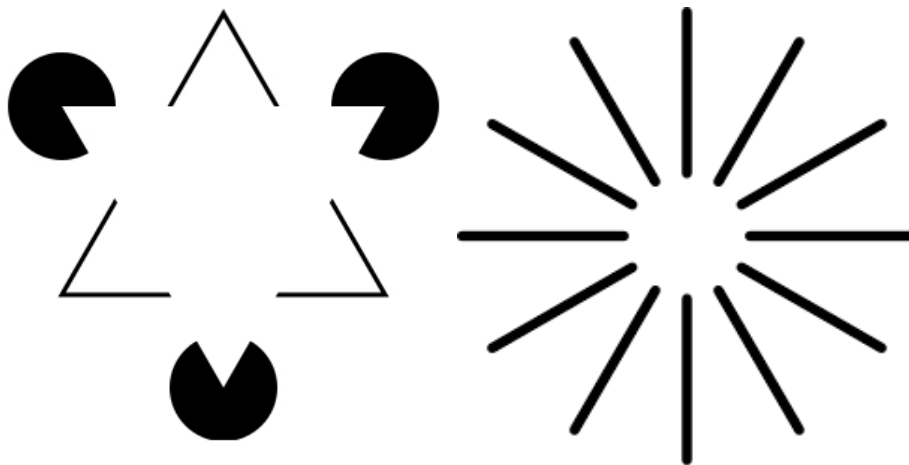
phenomena and other Gestalt properties. But these aspects are different from the aspect central to the recognition of something as a picture.

<sup>28</sup> Or at least almost constant, given that what changes is whether the awareness in question serves as the base for a dependent second kind of awareness or not.

whether it depicts a round or an elliptical object). Similarly, we may fail to notice other aspects; we may hallucinate them; and we may err about their nature.

Finally, just as it is generally not up to us which aspect we see a given object as having, it is generally not up to us which surfaces we recognise as pictures, and what we thereby see in them. True, we may be able to stop seeing a picture as a picture (e.g. by attending to it in a certain way); and we may have some control over whether we see the duck-rabbit drawing as a depiction of a duck or as a depiction of a rabbit. But we cannot bring ourselves to see a surface as a picture if the surface is not such as to give rise to seeing-in in normal viewers under suitable conditions, (e.g. a random wall of a building); and we cannot see a horse in the duck-rabbit drawing. Indeed, once we have recognised a flat surface as a picture, we may not be possible for us anymore to switch back to seeing it as a mere surface. Again, the same applies to other aspects that we see objects as having. As noted earlier, this lack of subjection to the will is part of the reason why both seeing-in and seeing an aspect should not be taken to involve imaginative awareness.

25. The Aspect View can also answer both the Question of Content and the Question of Depth. The answer to the first question is roughly the same as the one given by the Experienced Resemblance View. The reply to the second question points to the fact that the aspect seen is three-dimensional. But, of course, we do not really see depth (there is none); nor do we come to falsely believe that there is a three-dimensional scene before us, or that the picture does possess the visual appearance of such a scene. Rather, our visual awareness of the three-dimensional arrangement of objects is an imperfect visual illusion. Imperfect illusions are illusions that we are able to introspectively distinguish from veridical perceptions and thus normally do not take at face value (e.g. by endorsing them in belief or action). Seeing-in involves such an imperfect illusion because we can tell from the inside that it differs from seeing a three-dimensional scene face-to-face, especially with respect to the dimension of depth.



FIGURES 12 AND 13. Unknown Artists, *Kanizsa's Triangle & Ehrenstein Illusion*.

We may still sometimes — particularly when concerned with very realistic depictions that only just fall short of being a *trompe-l'oeil* — feel an inclination to judge that there is a three-dimensional scene before us. And this might perhaps explain why some people have thought that pictorial experience involves, or comes very close to, a perceptual illusion of whatever is depicted (e.g. Gombrich (1960 pp. 3ff.)). But, unless we are concerned with *trompe-l'oeils*, such inclinations are typically undermined or outweighed by our recognition of the illusory character of our awareness of depth and volume.

Nor is it more generally unusual that aspect perception is illusory with respect to the experienced aspect (which is why it may not always be appropriate to speak of *seeing* an aspect). Contour illusions like the Kanizsa and Ehrenstein illusions are good examples [FIGURES 12 AND 13]. We see the incomplete black circles as being grouped together in such a way as to form a white triangle, and the thick black lines as being organised in such a way as to circumscribe a white circle. But there is really neither a white triangle, nor a white circle before us. Hence, our perceptual experiences of these shape aspects are illusory. The same is true in the case of *seeing-in*, only that here the illusion is more widespread and concerns most aspects of the visual appearance of the three-dimensional scene in question, notably its dimension of depth.

26. Finally, the Aspect View has the resources to capture the characteristic features of *seeing-in*. First, the kind of aspect perception under consideration involves the visual awareness of the picture's surface and its visible features. Second, it also makes us aware of the three-dimensional scene that is depicted. In particular, this awareness is visual in the same way as many other experiences of aspects. Third, our awareness — and probably also our visual attention — is always divided between the surface and the aspect, given that we cannot stop being aware of one of them without also ceasing to have an experience of *seeing-in*. Fourth, our awareness of the surface and our awareness of the aspect are not independent of each other, but form a single, unified experience. In particular, we cannot experience the aspect without also seeing the surface; and our experience of the surface differs introspectively relative to whether we notice the aspect. Fifth, as already mentioned, experiencing the aspect in question involves experiencing depth.

Furthermore, the Aspect View is unitary in that it claims that *seeing-in* has a single, but complex content, according to which the visible features of the marks on the surface and the visual appearance of the depicted scene as standing in a certain relation. This relation — which I have described, in want of a better word, as a relation of organisation — is an instance or determination of the relation that, more generally, holds between any aspect and the relevant visible lower-level features.

Evidently, more needs to be said in elucidation and defense of the Aspect View.

But the preceding considerations should have made clear that this view has at least one advantage over its unitary rivals: it can accommodate both the experience of resemblance in two-dimensional shape and the experience of the third dimension of depth by understanding seeing-in in terms of aspect perception.<sup>29</sup>

## References

- Budd, Malcolm (1987). ‘Wittgenstein on Seeing Aspects’. In: *Mind* 96.381, pp. 1–17.
- (1991). *Wittgenstein’s Philosophy of Psychology*. London: Routledge.
- (1991/2008). ‘The Look of a Picture’. In: *Aesthetic Essays*. University College London, pp. 216–238.
- (1992). ‘Review of ‘Mimesis as Make-Believe: On the Foundations of the Representational Arts’ by Kendall L. Walton’. In: *Mind* 101, pp. 195–198.
- (1992/2008). ‘On Looking At a Picture’. In: *Aesthetic Essays*. Oxford: Oxford University Press, pp. 185–215.
- Dorsch, Fabian (2012). *The Unity of Imagining*. Berlin: De Gruyter, p. 485.
- (2015a). ‘Depiction and Imagination’. Under Review.
- (2015b). ‘Knowledge by Visualisation’. Under review.
- Gombrich, Ernst H. (1960). *Art and Illusion: A Study in the Psychology of Pictorial Representation*. London: Phaidon.
- Hopkins, Robert (1998). *Picture, Image and Experience*. Cambridge: Cambridge University Press.
- (2003). ‘What Makes Representational Painting Truly Visual?’ In: *Aristotelian Society Supplementary Volume*. Vol. 77. 1. Wiley Online Library, pp. 149–167.
- (2004). ‘Painting, Sculpture, Sight, and Touch’. In: *The British Journal of Aesthetics* 44.2, pp. 149–166.
- (2010a). ‘Inflected Pictorial Experience: Its Treatment and Significance’. In: *Philosophical Perspectives on Depiction*. Ed. by Catharine Abell and Katerina Bantinaki. Oxford: Oxford University Press, pp. 151–180.
- (2010b). ‘Sculpture and Perspective’. In: *The British Journal of Aesthetics* 50.4, pp. 357–373.
- Hyman, John (2003). ‘Subjectivism in the Theory of Pictorial Art’. In: *Monist* 86.4, pp. 676–701.
- (2006). *The Objective Eye: Color, Form, and Reality in the Theory of Art*. Chicago: University of Chicago Press.
- Kulvicki, John V. (2006). *On Images: Their Structure and Content*. Oxford University Press.

---

<sup>29</sup> I am very grateful to Malcolm Budd, with whom I had the great pleasure of discussing in much detail a previous version of this chapter, as well as to Robert Hopkins and the two editors of this volume. My research on this chapter was generously funded by the Swiss National Science Foundation (PP00P1\_139004) and the Fundación Séneca (18958/JLI/13).

- Lopes, Dominic (2005). *Sight and Sensibility: Evaluating Pictures*. Oxford: Oxford University Press.
- Martin, M. G. F. (2012). ‘Sounds and Images’. In: *The British Journal of Aesthetics* 52.4, pp. 331–351.
- McDowell, John (1984/1998). ‘Values and Secondary Qualities’. In: *Mind, Value, and Reality*. Cambridge (Mass.): Harvard University Press, pp. 110–129.
- McGinn, Colin (2004). *Mindsight*. Cambridge (Mass.): Harvard University Press.
- Newall, Michael (2009). ‘Pictorial Experience and Seeing’. In: *The British Journal of Aesthetics* 49.2, pp. 129–141.
- O’Shaughnessy, Brian (2003). *Consciousness and the World*. Oxford: Oxford University Press.
- (2012). ‘Seeing an Aspect and Seeing under an Aspect’. In: *Wittgenstein and the Philosophy of Mind*. Ed. by Jonathan Ellis and Daniel Guevara. Oxford: Oxford University Press.
- Peacocke, Christopher (1987). ‘Depiction’. English. In: *The Philosophical Review* 96.3, pp. 383–410. ISSN: 00318108.
- Stokes, Dustin (2013). ‘Cognitive Penetrability of Perception’. In: *Philosophy Compass* 8.7, pp. 646–663.
- Voltolini, Alberto (2012). ‘Towards a Syncretistic Theory of Depiction’. In: *Perceptual Illusions. Philosophical and Psychological Essays*. Ed. by Clotilde Calabi. London: Palgrave Macmillan.
- (2015). *A Syncretistic Theory of Depiction*. London: Palgrave Macmillan.
- Walton, Kendall (1990). *Mimesis as Make-Believe*. Harvard: Harvard University Press.
- (1992/2008). ‘Seeing-In and Seeing Fictionally’. In: *Marvelous Images: On Values and the Arts*. Oxford: Oxford University Press, pp. 133–142.
- (2002). ‘Depiction, Perception, and Imagination: Responses to Richard Wollheim’. In: *The Journal of Aesthetics and Art Criticism* 60.1, pp. 27–35.
- Wittgenstein, Ludwig (1953/2009). *Philosophical Investigations*. Ed. by P. M. S. Hacker and Joachim Schulte. 3rd edition. Oxford: Wiley-Blackwell.
- (1958). *The Blue and Brown Books*. Oxford: Blackwell.
- Wollheim, Richard (1965/1974). ‘On Drawing an Object’. In: *On Art and the Mind*. Cambridge (Mass.): Harvard University Press.
- (1980). *Art and its Objects: with Six Supplementary Essays*. 2d ed. Cambridge: Cambridge University Press.
- (1986). ‘Imagination and Pictorial Understanding’. In: *Proceedings of the Aristotelian Society, Supplementary Volumes* 60, pp. 45–60.
- (1987). *Painting as an Art*. London: Thames & Hudson.
- (1998). ‘On Pictorial Representation’. In: *The Journal of Aesthetics and Art Criticism* 56.3, pp. 217–226.

## Credits

With the exception of Figures 4 and 9, all images have been made available by Wikimedia Commons and are included under the Creative Commons Public Domain Mark 1.0 licence. Figure 4 is from wikiHow and is used under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 Generic licence. Figure 9 was originally published in Life Magazine, volume 58, number 7, on the 19th of February in 1965.