ESTETIKA
THE EUROPEAN JOURNAL OF AESTHETICS

Seeing in VR, without Seeing-In

CRITICAL NOTE

Luca Marchetti
University of Genoa, IT

HUP HELSINKI UNIVERSITY PRESS

FACULTY OF ARTS
Charles University

ABSTRACT

In *The Aesthetics of Virtual Reality* (2022), Grant Tavinor claims that VR is a technologically fancy kind of picturing and, more specifically, that VR headsets elicit proper seeing-in experiences. According to Tavinor, seeing a virtual environment through a stereoscopic headset elicits the same twofold experience as ordinary pictures: users simultaneously perceive the three-dimensional depicted scene – the virtual environment – and the bidimensional surface responsible for displaying such a scene. In this critical note, I argue that this is a wrong characterization of the visual experience VR stereoscopic headsets elicit. I maintain that VR visual experiences are not twofold but onefold. When using a VR headset, we do not perceive the configuration as an organizational structure; our visual system does not represent, even unconsciously, the properties of the configuration as such. My argument proceeds in two directions: first through a phenomenological analysis of VR experience and, second, by corroborating these phenomenological observations with evidence from perceptual psychology. I conclude that, if VR environments are to be considered bona fide pictures, it is not because they elicit a twofold seeing-in experience.

CORRESPONDING AUTHOR:
Luca Marchetti
University of Genoa, IT
luca.marchetti@edu.unige.it

KEYWORDS:
virtual reality; seeing-in; pictorial experience; visual experience; perception; stereoscopy

TO CITE THIS ARTICLE:
Marchetti, Luca. 'Seeing in VR, without Seeing-In.' *Estetika: The European Journal of Aesthetics* LXII/XVIII, no. 1 (2025): pp. 71–81. DOI: https://doi.org/10.33134/eeja.534

DOI: 10.33134/eeja.534

Grant Tavinor's *The Aesthetics of Virtual Reality* is arguably the most comprehensive work on VR in the field of analytic aesthetics. The main claim Tavinor argues for in the book is that VR is a 'technologically fancy kind of picturing'[1] – that is, a new and peculiar kind of depictive technology. In this regard, after an inquiry into the concepts of medium and of virtuality (Chapter 2) and a thorough explanation of how VR media allow for a combination of virtual perspective and interactivity (Chapter 3), in Chapter 4, 'VR as a Picturing Medium', Tavinor takes the technical discussion of the previous chapter and situates it in the philosophical perspective of the ongoing debate about the nature of picturing and picture perception. In particular, Chapter 4 addresses some of the philosophical issues with seeing VR as a pictorial form and defends a twofold theory of seeing in VR. Seeing a virtual environment through a stereoscopic headset, Tavinor argues, elicits the same twofold experience as ordinary pictures: users of VR simultaneously perceive the three-dimensional depicted scene – the virtual environment – and the bidimensional surface responsible for displaying such a scene. For Tavinor, VR headsets elicit proper seeing-in experiences.

In this paper I argue, *pace* Tavinor, that this is a wrong characterization of the visual experience VR stereoscopic headsets elicit. VR visual experiences, I show, are not twofold but onefold. In fact, when using a VR headset, we do not perceive the configuration as an organizational structure; our visual system does not represent, even unconsciously, the properties of the configuration as such. My argument proceeds in two directions: first through a phenomenological analysis of VR experience and, second, by corroborating these phenomenological observations with evidence from perceptual psychology. I conclude that, if VR environments are to be considered bona fide pictures, it is not because they elicit a twofold seeing-in experience.[2]

In Tavinor's Section 4.3, 'VR and the Picturing Surface', where he defends a twofold theory of VR seeing, he writes that 'never in my experience of VR, was I not also aware that I was viewing images on a screen, and indeed that the features of this screen were perfectly evident if I attended to them' (*AVR*, p. 66). Later on, Tavinor makes an even stronger claim about the phenomenology of VR seeing, stating that 'users are usually somewhat aware of – if not paying particular attention to – the surface. The phenomenology of stereoscopic headset use, partly because of the infancy of the technology and its resulting perceptible artifacts, is ordinarily one of being aware of the screen' (*AVR*, p. 68). He then goes on, providing further justification for his twofold account of VR seeing, writing that 'it is clear, given the discussion of how three-dimensional scenes are perceptually derived from the dual images of the stereoscopic headset and the features of the visual encounter with these stereoscopic images [...], that users of VR do perceive the configuration of the pictorial surface' (*AVR*, p. 68).[3]

---

1 Grant Tavinor, *The Aesthetics of Virtual Reality* (London: Routledge, 2022), 12. Hereafter: *AVR*.

2 In another paper, I argue along similar lines against Voltolini's idea that we see-in mirrors. Luca Marchetti, 'Seeing in Mirrors, Without Seeing-in', *Philosophia* (2025). Alberto Voltolini, 'Seeing in Mirrors', *Journal of Aesthetics and Art Criticism* 79 (2021): 315–27.

3 Here, Tavinor specifies that 'perceive' means 'the visual selection, organization, and interpretation of the surface features, and the building of this into a visual scene' (ibid). This is different from mere sensation – simple retinal stimulation – and attention – consciously focusing on a visual stimulus. Tavinor offers a conceptual explanation of what it means to 'see' the screen in 'twofold seeing', which means not just sensing or attending to it, but truly 'perceiving' it. For Tavinor, this explanation, along with a discussion of the stereoscopic screen of a VR headset, should show that users genuinely 'perceive' the configurational aspects of the marked surface. Yet, if the argument of my paper holds, this is not the case.

DOI: 10.33134/eeja.534

Following Tavinor's lead, I will thoroughly analyse how three-dimensional scenes are perceptually derived from the dual images of the stereoscopic headset. I will also examine the features of the visual encounter with these stereoscopic images, arguing that these are exactly what distinguish VR seeing – which is onefold – from ordinary pictorial seeing, which is twofold. Rather than examining Tavinor's argument in defence of seeing-in VR, I provide a direct counterargument to his theory: if you accept my view on VR seeing, twofoldness is automatically ruled out as a correct characterization of VR visual experience.

Before proceeding to my main argument, let me make a remark on Tavinor's phenomenological considerations: it is worth noting that one may not necessarily share Tavinor's phenomenological reports. In fact, I do not agree with his description of VR experience. In contrast to Tavinor, never in my experience of VR was I aware that I was viewing images on a screen or that the features of this screen were evident when I focused on them. On the contrary, every VR experience I have had felt strikingly different from traditional pictures, mainly because I felt I was directly perceiving the virtual environment, not depicted objects on a surface as the twofold theory suggests. Furthermore, Tavinor does not cite other phenomenological reports besides his own. Referencing other sources would have strengthened his claim or – given that, to the best of my knowledge, no one has defended a twofold characterization of VR experiences – made his phenomenological claim more plausible. By contrast, the vast majority of VR scholars,[4] as well as VR headset engineers, creators of VR content, and gamers,[5] seem to think that in VR we do not see the screen but only the virtual environment. My introspective claim that we do not perceive the screen of VR as a configurational structure is not only mine; it is shared by other people as well. This provides a counterargument to Tavinor's phenomenological claim and a prima facie reason for rejecting twofoldness in VR.

I start by spelling out the basic consensus about what twofold experiences, like seeing ordinary pictures, are.[6] As is well known, pictorial experience is defined by

4 See Oliver Grau, *Virtual Art: From Illusion to Immersion* (Cambridge, MA: MIT Press, 2003); Paweł Grabarczyk and Marek Pokropski, 'Perception of Affordances and Experience of Presence in Virtual Reality', *Avant* 7 (2016): 25–44; Aubrey Anable, *Playing with Feelings: Video Games and Affect* (Minneapolis: University of Minnesota Press, 2018); David J. Chalmers, *Reality+: Virtual Worlds and the Problems of Philosophy* (New York: Penguin, 2022).

5 I provide two examples from statements by non-philosophers to support my claim: 'You can think of [VR content] as an image, but really it's just data for the brain' (Scott Broock, quoted in Carina Chocano, 'The Last Medium', *California Sunday Magazine*, October 2014, https://story.californiasunday.com/virtual-reality-hollywood/) and 'VR eliminates the need for external frames. For the first time, the medium is no longer outside us, but within us. The paint is human experience, and the canvas is our consciousness. The idea of an externalized medium ceases to exist. That's why I think of VR as the last medium' (Chris Milk, 'The Future of Virtual Reality', *Medium*, 16 June 2016, https://medium.com/@Within/welcome-to-within-c7d3daba2b55).

6 Apart from endorsing Bence Nanay's dorsal-ventral hypothesis of picture perception – outlined, for instance, Nanay's article '*Trompe l'oeil* and the Dorsal/Ventral Account of Picture Perception', *Review of Philosophy and Psychology* 6 (2014): 181–97 – Tavinor does not say much more about what seeing-in experiences are. To defend a twofold theory of our experience of an artefact, one should first have a clear understanding of the conditions for such an experience. Notably, Tavinor acknowledges that Nanay's theory may conflict with VR visual media: the dorsal stream cannot attend to the pictorial scene, meaning 'depicted objects are not represented in our egocentric space' (*AVR*, pp. 59–62). Yet VR pictures seem to place depicted objects in egocentric space, allowing users to feel a strong impulse to interact with them and a heightened sense of 'presence'. In Chapter 4, Tavinor revises Nanay's twofold theory to accommodate egocentric picturing and interactivity. As argued later in this note, these features stem from VR's onefold nature, making it unnecessary to invoke counterfactual relationships or redefine twofold experiences, as Tavinor does in Section 4.5.

DOI: 10.33134/eeja.534

Richard Wollheim as a 'seeing-in' experience – when we face a picture, we see the picture's subject *in* a marked surface.[7] And seeing-in experiences are experiences with a *sui generis* phenomenology, which he calls 'twofoldness'. Looking at a picture, a viewer undergoes a 'twofold' experience: on one hand, she is visually aware of the flat surface of the picture; on the other, she experiences the subject matter of the picture. Wollheim calls the first of these folds of seeing-in the 'configurational' fold and the second the 'recognitional' fold, and claims that these are two distinguishable but 'inseparable' aspects of a single visual experience, and not two experiences somehow combined.[8] From the standpoint of mainstream vision science[9] and our best philosophical account of pictorial experience,[10] the unique character of twofold experiences is understood at the level of visual representational content.[11] In cases of genuine twofold pictorial experiences, the visual system represents the scene and the surface as conjoint – this experience amounts to a fusion experience, where the two folds are interpenetrated.[12] As Wollheim says, neither fold coincides with the corresponding perceptual experience, either of the picture's vehicle or of the picture's scene, taken in isolation.[13] Without conjunction the experience is onefold and loses its unique 'pictorial' phenomenology.

In fact, not only is ordinary face-to-face seeing obviously onefold but some cases of picture perception are as well. On the one hand, we can have experiences of pure material surfaces without seeing a three-dimensional scene, as shown by aspect dawning figures before the dawning of aspects.[14] On the other hand, we can experience three-dimensional scenes without representing surface features. Briscoe, who discusses these cases thoroughly, provides two main examples: peephole seeing, 'viewing a single picture through a narrow aperture', and stereoscopic

7 Richard Wollheim, *Art and Its Objects*, 2nd ed. (Cambridge: Cambridge University Press, 1980), 137–51.

8 See Alberto Voltolini, 'Different Kinds of Fusion Experiences', *Review of Philosophy and Psychology* 11 (2020): 203–22.

9 See Reinhard Niederée and Dieter Heyer, 'The Dual Nature of Picture Perception', in *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*, ed. Heinrich Hecht, Robert Schwartz, and Margaret Atherton (Cambridge, MA: MIT Press, 2003), 77–98; Rainer Mausfeld, 'Conjoint Representations and the Mental Capacity for Multiple Simultaneous Perspectives', in Hecht, Schwartz, and Atherton, *Looking into Pictures*, 17–60; Bill Millar, 'The Conflicted Character of Picture Perception', *Journal of Aesthetics and Art Criticism* 64 (2006): 471–77; Dhanraj Vishwanath, 'Toward a New Theory of Stereopsis', *Psychological Review* 121 (2014): 151–78.

10 Robert Briscoe, 'Depiction, Pictorial Experience, and Vision Science', *Philosophical Topics* 44 (2016): 43–82; 'Gombrich and the Duck-Rabbit', in *Aspect Perception after Wittgenstein: Seeing-As and Novelty*, ed. Michael Beaney, Brendan Harrington, and Dominic Shaw (New York: Routledge, 2018), 49–88; Nanay, '*Trompe l'oeil*'; Alberto Voltolini, *A Syncretistic Theory of Depiction* (New York: Palgrave Macmillan, 2015); Voltolini, 'Different Kinds'; John Zeimbekis, 'Seeing, Visualizing, and Believing: Pictures and Cognitive Penetration', in *The Cognitive Penetrability of Perception: New Philosophical Perspectives*, ed. John Zeimbekis and Athanassios Raftopoulos (Oxford: Oxford University Press, 2015), 75–99.

11 See also Briscoe, 'Depiction, Pictorial Experience', 44–45.

12 Voltolini, 'Different Kinds'.

13 Richard Wollheim, *Painting as an Art* (Princeton: Princeton University Press, 1987), 46. For discussion of 'conjoint' representation of pictorial space and the pictorial surface, see Mausfeld, 'Conjoint Representations'; Millar, 'Conflicted Character'; and Briscoe, 'Gombrich and the Duck-Rabbit'. Nanay's theory, on which Tavinor builds his account, explains this interpenetration in terms of interaction of the two visual subsystems. See Nanay, '*Trompe l'oeil*'.

14 Voltolini, *Syncretistic Theory of Depiction*, 81–82.

DOI: 10.33134/eeja.534

imaging, 'viewing [...] two copies of the same picture through a stereoscope'.[15] In these cases, the display conditions – united, in stereoscopes, with the peculiar nature of the stereoscopic display – effectively eliminate cues that specify the presence and properties of the pictorial surface while leaving accessible sources of optical information that specify depth and 3D structure in pictorial space. These cases show that a patterned surface will induce the experience of pictorial space if it reflects or emits light that would elicit a 3D-scene-representing visual experience when reflected or emitted to the eye from the real world. However, it is not necessary for the surface to also reflect or emit light enabling perception of the surface's own properties, such as its absolute distance from the perceiver's eyes, orientation, gloss, or texture. When this happens – that is, when the visual system does not have access to surface features – pictorial experience is onefold rather than twofold.[16] This also occurs in other examples of pictures that do not elicit twofold seeing, like stereograms and autostereograms, 3D movies, large images seen at a distance, and successful instances of illusionistic painting (that is, genuine *trompe l'oeils*).[17]

This discussion has already highlighted how and why Tavinor is wrong in characterizing stereoscopic imaging as a twofold experience: stereoscopes do not trigger 'conjunction' and VR headset technology is based on stereoscopic picturing. First, in stereoscopic imaging, the depicted objects are perceived as detached from the surface that hosts them due to parallax effects elicited by stereoscopy, that is, the separation of the right and left images on the screen.[18] In negative parallax stereoscopy, as seen in 3D cinema, depicted objects appear in front of the screen: we perceive a protruding object that is detached from the surface that produces it and 'invades' the space of the spectator. In positive parallax stereoscopy, as seen in stereoscopes, depicted objects appear behind the screen: we perceive a virtual scene behind a glassy surface, and the two bear no significant perceptual relation.[19] At most, I claim, in stereoscopes we *see through* but do not see in.

---

15 Briscoe, 'Depiction, Pictorial Experience', 58. A stereoscope is a device for viewing a stereoscopic pair of separate images, depicting left-eye and right-eye views of the same scene, as a single three-dimensional image. VR headset technology is based on stereoscopic picturing. For a thorough and technical discussion of the technology behind VR imaging see *AVR*, pp. 51–54.

16 In Nanay's terms, this could be explained by the fact that the dorsal stream is not representing relevant surface features, but only features of the depicted scene – as Gabriele Ferretti has convincingly argued regarding genuine *trompe l'oeils*; see his 'Why *Trompe l'oeils* Deceive Our Visual Experience', *Journal of Aesthetics and Art Criticism* 78 (2020): 33–42.

17 While peephole seeing and genuine *trompe l'oeils* have been much discussed in the literature, stereoscopic imaging has been curiously neglected. This is unfortunate because stereoscopic imaging is a common type of imaging – 3D cinema has been around, with varying levels of success, for about 70 years – and it is likely to become more common in the future, as the VR case makes plainly evident. The discussion that follows, apart from providing a counterargument to Tavinor's twofold account of VR seeing, can also be seen as a first attempt to qualify a feature of our visual experience of stereoscopic imaging, its onefold nature.

18 Your eyes are approximately two and a half inches or six centimetres apart ('interocular distance'), so they see the same image from slightly different angles and perspectives. Your brain then combines these two images to gauge distance. This is called binocular vision. Stereo vision, or 'stereopsis', is a result of good binocular vision, wherein the separate images from two eyes are successfully combined into one 3D image in the brain. When shooting a 3D image, two cameras are used to capture separate images of the same object from slightly different angles at one fixed viewpoint. When played back on a plano-stereoscopic display, the left image is shown only to your left eye and the right image only to your right eye. Your brain then fuses these two images to give you a perception of depth.

19 Simon Reeve and Jason Flock, 'Basic Principles of Stereoscopic 3D', *Sky UK*, 2010, https://www.ncl.ac.uk/media/wwwnclacuk/pressoffice/files/pressreleaseslegacy/Basic_Principles_of_Stereoscopic_3D_v1.pdf.

DOI: 10.33134/eeja.534

Second, VR technology, as a refined kind of stereoscopic imaging, enforces positive parallax with further depth cues that enhance the sense of depth independently from the surface hosting the scene. Additionally, VR obliterates the perception of the surface entirely using unique displaying conditions. VR headsets, with their specific stereoscopic display and unique displaying conditions, do not elicit twofold perception – there is no conjoint representation of a 3D scene and a 2D surface. VR visual experience should be seen as a case of onefold experience.

In the remainder of this paper, I clarify and defend these claims. I identify three main phenomenological contrasts that differentiate twofold pictorial experience from onefold seeing and show that these same differences exist between twofold pictorial experiences and VR experiences.[20] Additionally, I explain these phenomenological differences through their psychological underpinnings, supporting my phenomenological claims with data from perceptual psychology. Essentially, the three contrasts will allow me to show, first, that when using a VR headset the user's visual system cannot access information from surface properties as organizational elements of a three-dimensional scene – Contrast (1) – and, second, that absolute depth cues, which in the case of twofold experiences are triggered by the surface and used by the visual system to represent the distance of the surface from the viewer, in VR are triggered by the virtual environment and used by the visual system to represent the distance between the viewer and the virtual objects – Contrasts (2) and (3). Taken together, these discussions provide a direct argument for a onefold account of VR visual experience and a strong counterargument to Tavinor's twofold account.

Contrast (1). In twofold pictorial experiences, 'we are visually aware of at least some of the properties exemplified by the 2D, pictorial surface'; that is:

> it may not be psychologically possible to experience the same solid angle in the visual field as filled by a nontransparent, 2D surface at a single distance in depth and by an array of 3D objects located at different distances in depth simultaneously. But plausibly it is possible, when looking at a picture, to divide attention between the region of phenomenally 3D pictorial space contained within some solid visual angle θ and the distribution of 2D pictorial surface properties contained within some different solid visual angle j.[21]

This is something possible only in twofold experiences, and impossible in ordinary face-to-face seeing. However, I argue that it is equally unattainable in VR seeing – or in stereoscopic imaging more generally. In fact, no matter the effort, it is impossible to view the VR surface as a two-dimensional layout that the visual system arranges in specific ways to project a three-dimensional scene beyond it. In simpler terms, objects seen in VR – or in stereoscopic imaging more generally – do not seem to have any significant link to the surface hosting them. As we have seen above, in positive parallax stereoscopy as realized in stereoscopes, depicted objects appear behind the screen: we perceive a virtual scene behind a glassy surface, and the two bear no significant perceptual relation – there is no conjunction, and no perception that the virtual space is shaped by the surface enabling its perception. This is because of how our visual system processes stereoscopic images. First, the fusion of the two

20 I am following Briscoe, 'Depiction, Pictorial Experience'. While he discusses the contrasts between pictorial experience and face-to-face seeing, he never considers VR.

21 Ibid., 59–60.

DOI: 10.33134/eeja.534

stereoscopic images – the elaboration of binocular disparity by the visual system – automatically detaches the depicted objects from the surface that hosts them, positioning them inside the virtual space. This does not happen in ordinary two-dimensional pictorial experiences. Proof of the detachment of the scene from the surface – the perceptual dislocation of the former behind the latter – comes from a physiological response related to another depth cue, convergence. In stereoscopic imaging experiences, the eyes converge to the apparent position in the virtual space that the virtual objects occupy,[22] whereas in ordinary two-dimensional experiences our eyes converge towards the surface of the picture, not towards the depicted objects. This fact about how our visual system processes stereoscopic imaging, I claim, eliminates *conjunction* – the surface is not perceived as the relevant source of the three-dimensional scene we perceive. As a result, the stereoscopic experience is one-dimensional rather than two-dimensional. In stereoscopic pictures, the surface is not represented as such because all relevant depth cues are related to the virtual space and do not represent the surface.

VR, though, does even better than mere stereoscopic imaging. First, it integrates these already powerful cues with motion feedbacks, which enhances the detaching effect. Second, VR perceptually eliminates the screen – even as a mere interference surface – through peculiar displaying conditions. The headset's screen is positioned too close to the user's eyes (50–70 mm) to be clearly perceived as such. Moreover, in VR there is no framing. As Tavinor says, 'experiencing VR media through a stereoscopic headset certainly does not feel much like viewing a picture "from the outside" because the depicted scene is all around you' (*AVR*, p. 68). This is due to the display's width and the motion feedback information; that is, the visual perspective on the scene changes as the user moves their head, similar to face-to-face vision (more on this later). While I think that successful VR definitively eliminates the screen in this sense, you do not have to accept this stronger view. It suffices that something is seen through something else for this experience to be a mere onefold experience and not a twofold one. My analysis of the first feature of stereoscopic imaging already supported that conclusion: in VR, there is no representation of surface features in the 'conjoint' sense relevant for identifying a two-dimensional experience. I will now provide further evidence to support my claim with two more contrast cases.

Contrast (2). 'Experiences of pictorial space convey rich information about relative distance in depth (object a in pictorial space, for example, may be experienced as about twice as close to the pictorial point of view as object b), but they do not typically represent their objects as located at certain absolutely scaled distances in depth.'[23] Pictorial experience, unlike ordinary visual experience, is highly indeterminate with respect to the egocentric, viewer-relative locations of its objects. Consequently, it does not typically give rise to a robust impression of stereopsis, that is, solid appearance and immersive space. VR, on the other hand, does. This contrast is explained by a distinction Vishwanath draws between experiencing depth and 3D structure and experiencing *stereopsis* – the experience of stereopsis is characterized as the 'vivid impression of tangible solid form, immersive negative space, and realness that

22 Tavinor also acknowledges and considers this an 'illusionistic' merit of VR (*AVR*, pp. 69–71). In that passage, he seems to endorse an illusionistic, onefold view of VR experience. The problem is that he believes it can align with a twofold account of picture perception. However, since the two theories contradict each other, it is hard to see how this alignment is possible.

23 Briscoe, 'Depiction, Pictorial Experience', 60.

DOI: 10.33134/eeja.534

occurs under certain viewing and stimulus conditions'.[24] Pictures, real-world scenes, and VR all elicit the former experience – that of depth and 3D structure. But ordinary pictures, when the experience is twofold, *only* elicit the former experience without eliciting the latter: no *stereopsis*.[25] Real-world scenes and VR, on the other hand, both also elicit *stereopsis*. That is because VR engages the depth perception cues for the virtual environment in the same way as the real world does. This includes not only monocular cues – also called pictorial cues – such as occlusion, relative sizing, texture gradients, shadowing, and linear and atmospheric perspective, but also a range of powerful additional cues such as binocular disparity and convergence (as previously discussed), and kinetic cues such as motion parallax and da Vinci stereopsis. The important point I want to make here is not merely that the engagement of depth cues allows for the construction of a richer depth of the virtual scene and for a more volumetric perception of depicted objects – even though this is indeed an important feature of the perceptual character of VR, as Tavinor also acknowledges (*AVR*, pp. 51–54). Rather, the crucial point is that this engagement shows how the surface is not represented as such; it is not perceived as conjoint with the scene.

To understand why, we must introduce Vishwanath's absolute depth scaling hypothesis (ADSH) and distinguish between relative depth and absolute distance information in the visual system. Relative depth cues – such as occlusion, texture gradients, shading, kinetic depth, height in the visual field, and binocular disparity – indicate depth ratios (for example, object *A* is twice as far away as object *B*) and ordinal relations (for example, *A* is behind *B* but in front of *C*) but do not specify absolute distances. Absolute distance cues, like motion parallax, defocus blur, and oculomotor signals (such as convergence and accommodation), provide metric distance information. To represent a 3D scene, relative cues must be scaled by absolute cues. According to Vishwanath's ADSH, stereopsis arises in real-world binocular viewing because both types of cues are typically available. In pictorial space, however, only relative depth cues are present, explaining the absence of stereopsis:[26]

> When a picture is viewed binocularly, distance cues such as binocular convergence, vertical disparity and the accommodative state of the lens specify the distance of the visible picture surface, so there are no optical distance cues that specify the distance of pictorial objects. Pictorial depth cues, such as shading, perspective, and interposition, can specify the 3-D shape and relative layout of objects in the depicted scene, but without distance information, these cues cannot be scaled to derive absolute depth or size.[27]

---

24 Vishwanath, 'Toward a New Theory of Stereopsis', 153. See also Briscoe, 'Depiction, Pictorial Experience', 51–61; 'Gombrich and the Duck–Rabbit', 77.

25 For discussion of the absence of absolutely scaled depth information and stereopsis in pictorial experience, see, Dhanraj Vishwanath and Paul B. Hibbard, 'Seeing in 3-D with Just One Eye: Stereopsis without Binocular Vision', *Psychological Science* 24 (2013); Vishwanath, 'Toward a New Theory'; Briscoe, 'Depiction, Pictorial Experience', 61. Zeimbekis ('Seeing, Visualizing, and Believing', 321–23) also notes that the lack of stereopsis and parallax underlies the phenomenal difference between twofold picture perception and ordinary object perception or onefold experiences such as genuine *trompe l'oeils*.

26 For a more detailed discussion see Briscoe, 'Depiction, Pictorial Experience', 57n8, 61; 'Gombrich and the Duck–Rabbit', 77–78.

27 Vishwanath, 'Toward a New Theory', 158–59.

DOI: 10.33134/eeja.534

Ordinary pictures elicit two-dimensional experiences also because absolute cues for depth specify the distance between the viewer and the surface of the picture, not between the viewer and the depicted scene. In VR, that is not the case: both sources of relative depth and the main absolute distance information specify spatial features of the virtual environment.[28] This is because the surface is not represented as such. As discussed in connection with the ADSH, if the surface were represented, the visual system would not be able to process the scene as it does. If absolute depth cues represented the surface of the headset, we would be incapable of egocentrically locating the objects in the virtual space. But VR does the opposite because surface features are not represented. This is further evidence that in VR the visual system does not conjointly represent surface and scene features but only and consistently the features of the scene. Seeing in VR is onefold, not twofold. Contrast (3) further strengthens the point and provides my final evidence for VR as a onefold experience.

Contrast (3). 'When a subject moves in relation to a stationary, real-world object, the object's orientation doesn't visually appear to change. By contrast, when the subject moves in relation to certain pictures, for example, the famous British army recruiting poster depicting Lord Kitchener, the object she experiences in pictorial space may curiously appear to rotate toward her.'[29] In this respect, VR experiences are comparable to face-to-face vision. The reason lies, once again, in the fact that in VR there is no conjoint visual representation of the pictorial scene and surface: only the depicted content is represented, not the surface. It is the depicted content that changes as we move, rather than the viewpoint from which we see the surface. As previously mentioned, the VR surface may not be a fully encompassing visual field, but in VR there is no clear and perceivable framing. This is mostly because changes and movements of the viewer result in changes in the perspective of the depicted scene, not on the marked surface. After all, the surface is incredibly close to our eyes and moves with us as we move. Unlike ordinary pictures, changes in head movement result in changes in perspective on the virtual environment, not on the surface of the picture. This provides yet another reason to deny the conjoint representation of surface and scene in VR, which a twofoldness account would imply. VR experiences are onefold, not twofold as Tavinor maintains.

ACKNOWLEDGEMENTS

Thanks to the entire PEA group – Enrico Terrone, Nick Young, Vincenzo Grasso, Irene Olivero, Camilla Palazzolo, Gaia Penna, and Orsola Stancampiano – for helpful comments on an earlier version of this paper.

28 That is why, as Tavinor himself acknowledges, VR experiences result in the experience of a virtual environment with solid appearance and immersive space, one that is determinate with respect to the egocentric, viewer-relative locations of its objects in a way that twofold pictorial experience is not (*AVR*, pp. 72–83). Note that Tavinor accommodates egocentricity in his twofold account by introducing the idea that the central trick of VR picturing is to allow a counterfactual relationship between the apparent position of the scene viewer and the actual position of the picture viewer. But this claim is independent from the twofold theory: the counterfactual relationship is not between the actual viewer and the surface of the picture – whose position remains always the same, since it is fixed in front of the user's eyes – but between the actual position of the viewer and the perspective on the virtual environment, which is all on the side of depicted content. An account of VR egocentricity – even Tavinor's – depends only on one fold of the experience, the fold representing the virtual environment.

29 Briscoe, 'Depiction, Pictorial Experience', 60.

FUNDING INFORMATION

This research was funded by the ERC-StG Project 101040535 ('PEA – The Philosophy of Experiential Artifacts'), funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

COMPETING INTERESTS

The author has no competing interests to declare.

REFERENCES

Anable, Aubrey. *Playing with Feelings: Video Games and Affect*. Minneapolis: University of Minnesota Press, 2018. DOI: https://doi.org/10.5749/j.ctt20mvgwg

Briscoe, Robert. 'Depiction, Pictorial Experience, and Vision Science.' *Philosophical Topics* 44 (2016): 43–82. DOI: https://doi.org/10.5840/philtopics201644217

––––––. 'Gombrich and the Duck–Rabbit.' In *Aspect Perception after Wittgenstein: Seeing-As and Novelty*, edited by Michael Beaney, Brendan Harrington, and Dominic Shaw, 49–88. New York: Routledge, 2018. DOI: https://doi.org/10.4324/9781315732855-3

Chalmers, David J. *Reality+: Virtual Worlds and the Problems of Philosophy*. New York: Penguin, 2022.

Chocano, Carina. 'The Last Medium.' *California Sunday Magazine*, October 2014. https://story.californiasunday.com/virtual-reality-hollywood/.

Ferretti, Gabriele. 'Why *Trompe l'oeils* Deceive Our Visual Experience.' *Journal of Aesthetics and Art Criticism* 78 (2020): 33–42. DOI: https://doi.org/10.1111/jaac.12688

Grabarczyk, Paweł, and Marek Pokropski. 'Perception of Affordances and Experience of Presence in Virtual Reality.' *Avant* 7 (2016): 25–44. DOI: https://doi.org/10.26913/70202016.0112.0002

Grau, Oliver. *Virtual Art: From Illusion to Immersion*. Cambridge, MA: MIT Press, 2003. DOI: https://doi.org/10.7551/mitpress/7104.001.0001

Hecht, Heinrich, Robert Schwartz, and Margaret Atherton, eds. *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*. Cambridge, MA: MIT Press, 2003. DOI: https://doi.org/10.7551/mitpress/4337.001.0001

Marchetti, Luca. 'Seeing in Mirrors, without Seeing-in.' *Philosophia* (2025). DOI: https://doi.org/10.1007/s11406-025-00821-0

Mausfeld, Rainer. 'Conjoint Representations and the Mental Capacity for Multiple Simultaneous Perspectives.' In Hecht, Schwartz, and Atherton, *Looking into Pictures*, 17–60. DOI: https://doi.org/10.7551/mitpress/4337.003.0005

Millar, Bill. 'The Conflicted Character of Picture Perception.' *Journal of Aesthetics and Art Criticism* 64 (2006): 471–77. DOI: https://doi.org/10.1111/j.1540-594X.2006.00225.x

Milk, Chris. 'The Future of Virtual Reality.' *Medium*, 16 June 2016. https://medium.com/@Within/welcome-to-within-c7d3daba2b55.

Nanay, Bence. '*Trompe l'oeil* and the Dorsal/Ventral Account of Picture Perception.' *Review of Philosophy and Psychology* 6 (2014): 181–97. DOI: https://doi.org/10.1007/s13164-014-0219-y

Niederée, Reinhard, and Dieter Heyer. 'The Dual Nature of Picture Perception.' In Hecht, Schwartz, and Atherton, *Looking into Pictures*, 77–98.

Reeve, Simon, and Jason Flock. 'Basic Principles of Stereoscopic 3D.' *Sky UK*, 2010, https://www.ncl.ac.uk/media/wwwnclacuk/pressoffice/files/pressreleaseslegacy/Basic_Principles_of_Stereoscopic_3D_v1.pdf.'Basic Principles of Stereoscopic 3D.' *BSKYB*, 2010.

Tavinor, Grant. *The Aesthetics of Virtual Reality*. London: Routledge, 2022. Abbreviated as *AVR*. DOI: https://doi.org/10.4324/9781003107644

Vishwanath, Dhanraj. 'Toward a New Theory of Stereopsis.' *Psychological Review* 121 (2014): 151–78. DOI: https://doi.org/10.1037/a0035233

------, and Paul B. Hibbard. 'Seeing in 3-D with Just One Eye: Stereopsis without Binocular Vision.' *Psychological Science* 24 (2013): 1673–85. DOI: https://doi.org/10.1177/0956797613477867

Voltolini, Alberto. 'Different Kinds of Fusion Experiences.' *Review of Philosophy and Psychology* 11 (2020): 203–22. DOI: https://doi.org/10.1007/s13164-019-00456-7

------. 'Seeing in Mirrors.' *Journal of Aesthetics and Art Criticism* 79 (2021): 315–27. DOI: https://doi.org/10.1093/jaac/kpab021

------. *A Syncretistic Theory of Depiction*. New York: Palgrave Macmillan, 2015.

Wollheim, Richard. *Art and Its Objects*. 2nd ed. Cambridge: Cambridge University Press, 1980.

------. *Painting as an Art*. Princeton, NJ: Princeton University Press, 1987.

Zeimbekis, John. 'Seeing, Visualizing, and Believing: Pictures and Cognitive Penetration.' In *The Cognitive Penetrability of Perception: New Philosophical Perspectives*, edited by John Zeimbekis and Athanassios Raftopoulos, 75–99. Oxford: Oxford University Press, 2015. DOI: https://doi.org/10.1093/acprof:oso/9780198738916.001.0001

TO CITE THIS ARTICLE:
Marchetti, Luca. 'Seeing in VR, without Seeing-In.' *Estetika: The European Journal of Aesthetics* LXII/XVIII, no. 1 (2025): pp. 71–81. DOI: https://doi.org/10.33134/eeja.534

Submitted: 18 June 2024
Accepted: 31 January 2025
Published: 18 March 2025

Estetika: The European Journal of Aesthetics is a peer-reviewed open access journal published by Helsinki University Press in cooperation with the Faculty of Arts, Charles University in Prague.

HUP FACULTY OF ARTS Charles University