

## **Ksenija Puškarić**

Seton Hall University, 400 S Orange Ave, South Orange, NJ, U.S.A.  
ksenija.puskaric@gmail.com

# **The Argument from Hallucination**

### **Abstract**

*In this paper, I defend the argument from hallucination based on conceivability of Philosophers' hallucinations that involve no phenomenal difference with genuine perception. Essentially, this is a thought-experiment-based argument that defends the thesis that subjective indistinguishability of these two states implies the sameness of their objects. I address counterexamples and disjunctivist-style objections to this inference and expand the discussion toward conditions involved in conceiving philosopher's hallucinations. By sufficiently narrowing the "gap" between subjective indistinguishability and the sameness of objects thesis I counter Putnam's phenomenal-sorites-based objection and other related problems. Closer analysis of the conditions involved in the thought experiment shows that some commonly raised objections to the inference from the subjective indistinguishability thesis to the sameness of objects lose their plausibility.*

### **Keywords**

perception, hallucination, disjunctivism, sense datum, thought experiment, phenomenal sorites

## **Introduction**

Let me begin with a brief exposition of the background discussion on the philosophy of hallucinations. In all cases of hallucinations, one has a kind of perceptual experience; one *seems* to see something although there is no physical object being present to the visual senses. Such states are usually induced by drugs, alcohol, or abnormal neurological and psychological conditions. How do we account for the nature of hallucinations? One approach, like Robinson's (1994, p. 32), is to argue that while having a hallucination (H) one would be directly aware of a sense datum. Here is what I will call *the simple argument from hallucination*:

1. "If there sensibly appears to a subject to be something which possesses a particular sensible quality, then there is something of which the subject is aware which does possess that sensible quality." [Robinson's *Phenomenal Principle*]
2. If, for example, something red appears to me, then there *is* a direct object of my awareness which instantiates redness [from the *Phenomenal Principle* (PP)].
3. There is no physical object (no patch) which is red.
4. ∴ When hallucinating I am directly aware of a non-physical object (sense datum) which instantiates redness.

Let me explain this argument and its premises in more detail. Intuitively at least, it seems to be true that in order for redness to appear to someone, one has to be directly in contact with something red. For example, H. H. Price says that "[w]hen I say 'this table appears brown to me' it is quite plain

that I am acquainted with an actual instance of brownness” (1932, p. 63). On the other hand, A. D. Smith points out the following:

“To say simply that our subject is not aware of anything is surely to under-describe this situation dramatically. Perhaps we can make sense of there being ‘mock thoughts’, but can there really be such a thing as mock sensory awareness? Perhaps there can be ‘an illusion of understanding’, but can there be an illusion of awareness? [...] The sensory features of the situation need to be accounted for [...]. If we take as our example subjects who are fully attentive and focused, we need to do justice to the fact that such subjects in some sense take cognizance of, indeed fully attend to, sensory presentations. But if so, what else can we say other than that the subject is, as the Argument requires, aware of a non-normal object?” (Smith 2002, pp. 224–225)

If nothing red really strikes me perceptually, then what explains the quality of the experience? Since properties must be instantiated, and there is nothing red “outside”, it follows that the direct object of my awareness which instantiates redness must be a non-physical object, a sense datum. The so-called phenomenal principle (PP) formulated by Robinson has explanatory power. M. G. F. Martin furthermore explains the plausibility of the principle in the following way:

“While the principle is not self-evident, it is nonetheless *evident in the light of experience*. If all one has to go on is reflection on the proposition itself, then one cannot determine whether it is true or not. But the proposition in question concerns appearances, how things appear to one, and that one can test just by reflection on how things do appear to one. Hence, simple reflection on one’s own case should show one whether the principle in question is true.” (Martin 2004, p. 30)

If the PP is justified as a principle that accounts for the phenomenology of appearances, and in hallucinations something certainly appears such-and-such, then the PP must apply to hallucinations, too.

Let us suppose now that the argument so far holds for hallucinations (leaving aside the problems for the sense datum theory and the PP). The question now is how does the argument from hallucination generalize to the cases of veridical perception? This stage of the argument will be the focus of my paper. I am interested in exploring the philosophical significance of conceivable (i.e. possible) hallucinations which are for the subject phenomenologically exactly like veridical perception. Robinson calls them “philosophers’ hallucinations”:

“Philosophers are (mainly, at least) interested in what I shall call ‘philosophers’ hallucinations’. These are not, as far as we know, hallucinations as they actually occur, but they are, it is argued, the hallucinations that would occur if the perceptual system and brain were stimulated in the just the way it is stimulated in genuine perception, but directly and not by the usual external objects. This would give, it is supposed, a hallucination indistinguishable to the subject from the corresponding perception, which is not the case, at least in general, for hallucinations as they actually occur.” (Robinson 2013, p. 314)

Let me now explore the meaning of the “subjective indistinguishability” thesis.

### **Subjective Indistinguishability of Philosophers’ Hallucinations**

For example, you are now veridically perceiving letters on a piece of paper. But you might be having a philosopher’s hallucination; it is conceivable that you could be hallucinating reading these letters by being presented with an appearance which is phenomenologically just like the real object – subjectively indistinguishable from the appearance of real letters. A. D. Smith further clarifies:

“The restriction that such discrimination should be solely on the basis of the experience itself, or ‘what it is like’ to have the experience, is obviously required, since hallucinations can certainly be discriminated from perceptions in other ways. I might, for example, learn that I am hallucinating by being informed of the fact by someone who knows; or I might infer that I am hallucinating because I regard the apparently presented scene as physically impossible.” (Smith 2008, p. 182)

Let me begin with the thesis that it is possible that a hallucination is subjectively indistinguishable from VP. According to Byrne and Logue it could mean the following:

“We may define a case  $\alpha$  to be subjectively indistinguishable from the good case iff, in  $\alpha$ , the subject is not in a position to know by ‘introspection’ alone that he is not in the good case.” (Byrne & Logue 2009, p. 58)

In such a perfect hallucination, you would not be able to tell that you are not veridically seeing. Usually, things are subjectively indistinguishable in virtue of having something in common. For example, a real Easter egg and a plastic egg might easily be subjectively indistinguishable because they look very much alike. The “egg looking property” is something that both eggs share. Similarly, the question I want to raise now is in virtue of what shared property Hs and VPs are subjectively indistinguishable? The simplest and plausible explanation is that appearances share their properties. For example, seeing a red patch and hallucinating a red patch have something in common – a red looking property. This way, by introducing a “common factor”, it is possible to obtain a unified explanation of indistinguishable experiences. Thus, let me make a more general point: for every VP, it is conceivable to be in a mental state which is subjectively indistinguishable from a VP in virtue of its properties of appearance, although there is no physical object that one is directly aware of. If it is conceivable that H appearance is subjectively indistinguishable from V appearance, then it is possible that in both cases (properties of) appearances are the same. I proceed with the following formulation:

1. There are possible hallucinatory appearances that are subjectively indistinguishable from veridical appearances. [*The thesis about “philosophers’ hallucinations”*]
2. If hallucinatory appearances are subjectively indistinguishable from veridical appearances, then (properties of) appearances in both cases are the same.

I will refer to “(properties of) appearance” shared by H and VP because the term is theory neutral and non-committing; it screens how the purported object appears, “bracketing” the nature of the real object (if any). The 2<sup>nd</sup> premise figures as the best explanation of the possibility stated by the 1<sup>st</sup> premise. Since such a view is to great length discussed and challenged in contemporary philosophy of perception, I will introduce the main issues and positions involved in the debate.

### ***Considerations Based on the Subjective Indistinguishability Relation and the ‘Common Factor’***

Let me begin with an example. A tiger looks very much like a big cat because they both belong to the same family – cats. Being alike or being indistinguishable is usually explained by something that things share. Similarly, the indistinguishability of H and VP is explained in terms of their common kind. M. G. F. Martin says that:

“I shall call The Common Kind Assumption: that whatever kind of mental episode it is that is occurring when I am veridically perceiving – say when I am seeing the orange as it is – that same kind of episode can occur when I am merely having a hallucination, as when my optic nerves are suitably artificially stimulated. For such a view, although there can be all the difference in the world between a situation in which I am seeing an orange and one in which I am merely hallucinating one, there need not be a difference in the kind of experience or mental episode which occurs in both cases.” (Martin 2004, p. 7)

The so called “common factor view” (or “conjunctivism”) posits a common factor, a kind shared by an H and its corresponding VP which explains their subjective indistinguishability. Such a view is seriously questioned by the disjunctivists who (for one reason or another) deny that H and VP share the common kind.<sup>1</sup> So, let me introduce the main theses of disjunctivism:

*Disjunctivism – a version of direct realism – is the view according to which physical objects are direct objects of perception and essentially constitute experience.*

Consider an example introduced by Byrne and Logue:

“Imagine that you are looking at an ordinary lemon in good light. Your vision is good: you see the lemon, and it looks yellow and ovoid. Now suppose that, unbeknownst to you, some minor deity removes the lemon, while preserving its proximal neural effects. Your brain is in the same local physical states as it was in when the lemon was there: the neurons in your visual cortex, for instance, are firing in the same pattern. After the removal, you do not see the lemon, because the lemon is not around to be seen. Yet – we can all grant – you notice nothing amiss.” (Byrne and Logue 2009, p. vii)

As Johnston (2004) says, you have undergone a “subjectively seamless transition”. Now, the key question is whether your experience has changed after the lemon is removed? The disjunctivist says “yes” because the lemon is essential for the experience. The lemon “out there” constitutes your lemon-perception. Once the lemon is removed, there is no lemon-perception; there is an H of lemon. Your experience cannot be identified independently of its object. The common factor theorists argue the contrary – the experience did not essentially change. By removing the lemon, the object is changed, not the experience itself. What is essential for lemon-perception is not the lemon itself, but the perceptual experience or the mental kind shared by H and VP, which figures as the common factor and remains the same after the lemon is removed. They regard the subjective part, the internal component that did not change, to be essential for perception.

On the other hand, the disjunctivists say that H and VP just subjectively seem to be the same experience, but this criterion is not telling us about the real nature of those experiences. They reject the “common kind assumption”. According to McDowell, “an appearance that such-and-such is the case is either a mere appearance or the fact that such-and-such is the case making itself perceptually manifest to someone” (McDowell 1982, p. 472). Or, to use Neta’s (2008) example, the difference between H and VP is not like that one between two kinds of tigers, e.g. between a Bengal and Siberian one. Similarities between H and VP are like similarities between a real tiger and a hologram of a tiger. Yet they do have something in common, they both fall under a *disjunctive description*: either an experience is a VP, or it is a H. They also share the epistemic fact that they are indistinguishable from each other, from a certain vantage point. Thus, the common factor (in case of disjunctivism) is usually at best epistemic rather than metaphysical. However, nowadays there are various ways in which the common factor view can be spelled out:

for example, the common factor can be individuated in terms of sameness of their contents (Robinson, Crane 2005), and/or objects, appearances (Alston), introspectively detectable features (Neta). And the conjunctivist position can be sufficiently weakened to be compatible with some forms of disjunctivism. Although there are many arguments<sup>2</sup> challenging the inference from subjective indistinguishability to the common factor thesis, I will begin by focusing on one problem. Putnam (1999, p. 130) presented an argument against conjunctivism which is partially based on empirical considerations about subjective indistinguishability – the so-called “phenomenal sorites argument” (PSA).

### *Countering the Phenomenal Sorites Argument*

The subject of the phenomenal sorites experiment is presented with two blue cards, A and B, and she judges that they look the same. Thus, if *subjective indistinguishability implies sameness of appearances*, then  $A=B$  (“=” denotes “the sameness of appearances”, and “ $\approx$ ” denotes “subjective indistinguishability”). Then, the subject is presented again with two blue cards, B and C, and she judges that they look the same. Therefore, from the above implication it follows that  $B=C$ . From the transitivity of identity, it further follows that  $A=C$ . However, when presented with A and C cards, she judges that they *do not* look the same. But, if A and C are distinguishable, then the appearance of the A card is *not* identical to the appearance of the C card. Here is the paradox:

subjective indistinguishability ( $\approx$ )  $\rightarrow$  sameness of appearances<sup>3</sup> (=)

$(A \approx B) \rightarrow (A = B) \quad A = B$

$(B \approx C) \rightarrow (B = C) \quad B = C$

$\neg (A \approx C) \rightarrow \neg (A = C) \quad \therefore A = C$

$\approx$  is *not* transitive = *is* transitive

$\therefore \neg (\text{subjective indistinguishability} \rightarrow \text{sameness of appearances})$

On the left side of the implication, we are tracking subjective indistinguishability relation.

Although  $A \approx B$  and  $B \approx C$ , the relation of “subjective indistinguishability” fails to be transitive; it is not the case that  $A \approx C$ . Consequently,  $\neg (A = C)$ . In the experiment, appearances change gradually so that one cannot notice the change. Assuming that subjective indistinguishability implies sameness of appearance, and we trust our abilities to judge how things appear, we end up with a paradox;  $A = C$ , and yet  $\neg (A = C)$ . The phenomenal sorites shows that indistinguishability is *not* a transitive relation, although identity is transitive. Therefore, it cannot be true that subjective indistinguishability implies

<sup>1</sup> For more on this debate and relation to skepticism see (Farkas 2006), (Huemer 2000), (Huemer 2001), (Le Morvan 2004), and (Thompson 2008).

<sup>2</sup> For example, McDowell (2008) and Williamson (2000) argue that the argument relies on an erroneous theory of self-knowledge.

<sup>3</sup> Here one can substitute “sameness of appearance” with “sameness of objects” or “sameness of content”, to make the PSA argument directed against more versions of conjunctivism.

sameness of appearances. The PSA at least threatens some versions of conjunctivism, those that affirm that subjective indistinguishability implies sameness of objects (Robinson) or contents (Crane).

Now, the question is how strong the PSA-based objection really is. The PSA shows that our judgments about how things appear are not infallible in cases where appearances change gradually. However, it does not show that our judgments about how things appear are *generally unreliable*. To show such a strong point, namely, that we should not rely on our judgments, or that our discriminative powers are not reliable indicators of how appearances are, one would have to prove that we make mistakes *systematically*.

There is also another answer available to the conjunctivists: there is a relevant disanalogy between the phenomenal sorites case and Hs-VPs. Appearances in the PSA example change gradually; the change in blue color is too small to be detectable. But such a change is a matter of *degree*. On the other hand, if the disjunctivist is right, then Hs and VPs are different *kinds* of states – there is no common factor shared by H and VP. That difference might weaken the PSA argument as an objection. If we make mistakes when appearances differ in *degree*, that does not show that we would miss the difference between two *kinds* of appearances –hallucinatory and the veridical one. Let me give an example to make the point clearer. Suppose that I am giving a talk at a conference. Now, suppose that every five seconds of my veridical perception, an evil colleague induces my brain to have the same perceptual experience that I would have anyway; if they were not controlling my brain, I would have the same visual experience as I have when they induce it. Suppose that I have a VP for five seconds, and then I have a hallucinatory experience for five seconds, then again five seconds of a VP followed by five seconds of H. And yet, I do not notice any changes happening. My experience flows ‘smoothly’. Would it not be surprising if the kind of experience changed without my noticing any change? Neta argues that it is conceivable and possible not to notice a difference between two kinds of things, and he gives an example to illustrate the point:

“If each of those two things [real tiger and holographic tiger] is composed of great many parts then there could be a gradual transition from one thing to the other thing by a temporal series of replacements of parts of one for parts of the other. We could replace one small part of the tiger with one small part of the hologram, and then continue this series of replacements one-by-one, and thereby make a gradual transition from the tiger to a hologram of the tiger.” (Neta 2008, p. 138).

Although such an answer is plausible, my intuitions are more inclined toward the common factor view. Thus, I will explore the conditions of a thought experiment (TE) argument that involves imagining philosopher’s hallucinations, which could potentially meet the objection raised by the PSA and close the gap between subjective indistinguishability and sameness of appearances. Before presenting the argument, let me mention two more worries that are going to be addressed in the same manner.

The first question is how can one imagine a perfect hallucination indistinguishable from VP? Can our imagination be so powerful? Is it not possible that in our conceiving appearances merely at first seem to be the same? For example, if you try to close your eyes and imagine the perceptual scene that you have at this very moment, the imaginary scene certainly will not be as rich as the real perceptual scene. The worry is that in the argument from hallucination there are two appearances in different modes of presentation

– perception and imagination. Different modes normally make a difference in the phenomenology of the object presented, in vividness, colorfulness, etc. In that case, the same content (“what is experienced”) due to different modes of presentation should yield subjectively distinguishable states. Then, the objection is that an H imagined appearance in thought would not be phenomenally the same as perceived appearance. Consequently, we have good reasons to be skeptical of how the conceivability argument is supposed to work at all.

Another related worry is that two appearances may seem to be indistinguishable, and yet they might differ in how their objects really are. For example, a chiliagon (a polygon with one thousand sides) and a circle are indistinguishable although they are different figures. Similarly, it is possible that appearances in our thought are indistinguishable, even if a hallucinatory appearance in our conceiving may have something that the veridical appearance lacks (or vice versa), and we may not be able to spot a fine detail. Thus, the inference from subjective indistinguishability to sameness of appearances is thereby rendered questionable.

In what follows, I will attempt to address the two worries in more detail, and the objection based on the PSA.

### The Thought Experiment Argument

In this thought experiment (TE), you are asked to conceive a (hallucinatory) appearance which is indistinguishable from your present veridical experience of people sitting in front of you in a classroom, whereas there are really no people and no chairs in the external world. In Husserl’s (1931) terminology, one needs to use the so-called method of *epoché* or “bracketing”, according to which phenomenological specification of what the experience is about in the TE must not rely on the correctness of any existence assumption of the object (if any). The content of experience, in the present usage, captures the phenomenal aspect of the experience, and is identical with how the purported object is presented, “bracketing” the existence of its object (if any). An H appearance which is indistinguishable from a V one is an appearance which could be a V, as far as you can tell purely by reflection alone. So, what appearance do you imagine? The condition of the imagined scene is to imagine an appearance of people sitting in front of you, and add to it that “there are no physical people in front of me, no physical walls, no chairs, etc.”. However, this extra condition neither adds to nor in any way changes the act of imagining the scene you are presently aware of, because in the phenomenological description of a V appearance the existence of a physical object is “bracketed” in the first place. It is not the case that you have to visually imagine an extra tiny detail or make the appearance slightly blurry. So, the H appearance that you imagine remains the same as the veridical scene before your consciousness. In imagining a hallucinatory appearance which is indistinguishable from the V one you imagine your present veridical scene. But, to imagine a veridical scene is to keep the same apparent features before your mind. Nothing changes. Therefore, imagining an H indistinguishable from VP is to imagine an appearance which is *exactly like* a veridical appearance.

Let me present the TE reconstructed in the following form of an argument (“ $\approx$ ” denotes “subjective indistinguishability”, and “=” denotes “identical”):

1. To imagine an H appearance ( $\approx$  V appearance) is to imagine a H appearance such that it could be a V one, as far as you can tell by reflection alone.

2. To imagine an H appearance ( $\approx$  V appearance) is to imagine (a V appearance + “no real people in front of you, no real chairs, etc.”).
3. The “no real people in front of you, no real chairs, etc.” condition makes no difference to the imagined scene.
4.  $\therefore$  To imagine an H appearance ( $\approx$  V appearance) is to imagine an H appearance (=V appearance).
5.  $\therefore$  In the TE (or in imagining) there is no distinction between subjective indistinguishability and the sameness of appearance.

This argument shows that when you conceive an H appearance which is indistinguishable from a V you imagine an H appearance which is *the same* as a V appearance. However, the term “the same” is ambiguous here. Depending on interpreting the second step of the TE, it is possible to obtain two interpretations of the TE, and two corresponding conclusions:

- a. *The narrowest type identity*<sup>4</sup> of H and V: in the TE there are two appearances, the hallucinatory and the veridical one, which are exactly alike with respect to all reflectively accessible features.

According to option a. imagining (V appearance + “no real people in front of me”) happens in two stages. First, at  $t_1$ , you have a V appearance before your mind, and then at  $t_2$  you add the condition “no real people in front of me”. At  $t_2$ , you have imagined the required hallucination. If there are two mental episodes in the TE, then there are two appearances, H and V one. Therefore, a V appearance is qualitatively (or of the same narrowest type) identical with an H appearance. Consequently, the identity claim (=) in the fourth step of the TE amounts to the narrowest (or qualitative) type identity of an H and a V appearance. This would be one way to present how the skeptic can formulate the “matching” relation of veridical and non-veridical experience in the skeptical scenario.

- b. *Co-referentiality*: H and V appearance refer to the same purported object.

According to option b., imagining (V appearance + “no people in front of me”) happens in one mental act. You imagine something that you actually see, but what you add in imagining does not affect the scene that is present to your mind. Thus, the act of perceiving overlaps with the act of imagining both temporally and with respect to the purported object. Such an object, which must be the same in both cases, is apprehended in two modes: perceiving and imagining. The relation between H and VP with respect to their purported object is somewhat like Frege’s relation between the morning star and the evening star. The first star visible in the morning is the last star visible in the evening, and both aspects refer to the same thing. In the imagined reality, V and H appearances are co-referential; they share a purported object, while their difference is mode-individuated and remains unreflected in imagination.

The thought experiment argument renders the following two premises of the argument from hallucination:

1. It is conceivable (possible) that H appearance is subjectively indistinguishable from V appearance.
2. Therefore, appearances of H and VP are co-referential/narrowest-type identical.



The two different interpretations of the TE give two conclusions that differ in logical strength and have an impact on the three objections that motivated the introduction of the TE argument in the first place. Let me address these issues now.

The latter interpretation (b) of the TE directly addresses the first worry – that two modes of presentation must yield a difference on the phenomenal level. The answer is that in the TE you need not imagine an H appearance subjectively indistinguishable from a V appearance by imagining “something *like* V appearance, whereas there are no physical objects out of my mind”. This would imply that we have to imagine a “mental copy” of the V appearance. In this case it would be hard to see how you can imagine an appearance which is exactly like the V one, perfectly vivid in every detail. However, the skeptic need not require such a strong claim. Instead, they ask you to imagine “a V appearance, whereas there are no physical objects out of the mind”. Here, you do not need to make a “mental” copy but just add the condition “there are no physical objects out of the mind” to the veridical scene present to your consciousness. In this case, the imagination-mode does not make a difference on the phenomenal level to the purported object presented in the perceptual mode. Hence, the objection that imagination and perception (as two modes of presentation) would yield phenomenally different appearances can be answered.

Versions (a) and (b) also offer a solution to the problem raised by the PSA and the chiliagon counterexample. Both objections undermine a more general thesis, namely, that subjective indistinguishability of appearances does not imply sameness of appearances. However, the TE presents a special case for hallucinations. H and V appearances in the TE are either co-referential or the narrowest type identical. From this it follows that *it is impossible that there is a difference in detail between appearance of H and VP which is undetected*. If there is no chance, in principle, that appearances are subjectively indistinguishable and their appearances differ, then despite conceding the PSA objection and the counterexample with the chiliagon, the thesis that subjective indistinguishability implies sameness of appearances in cases of H-VP does not seem to be directly undermined.

The gap between subjective indistinguishability and sameness of appearance in the TE is (at least) sufficiently narrowed to meet the three objections. However, the first interpretation (a) is much weaker than (b), and it would not be sufficiently strong for the conclusion concerning the sameness of *objects* in the H and the VP case. The (a) strategy gives the conclusion that H and V appearances are narrowest-type identical. But, in this case, there will still be room for the possibility that qualitatively identical appearances are instantiated in different kinds of objects. For example, a holographic image of a tiger might look the same as a real tiger, and yet those objects are different. Similarly, it may be argued that in the case of an H, properties of appearance may be instantiated in a sense datum, but not in the case of VP.<sup>5</sup> Thus, the conjunctivist (sense-datum theorist) needs to offer an argument for the inference from the type identity to the object identity. Before I proceed with

4  
The term comes from M. G. F. Martin.

5  
This kind of “mixture” of disjunctivism and conjunctivism is defended by W. P. Alston.

that argument, let me address another objection<sup>6</sup> to conceiving Philosophers' hallucinations.

### *Disjunctivist's Objection from Conceivability*

Let me start with an example. If water is  $H_2O$  then it is not metaphysically possible that water is  $H_3O$ , and from this it follows that you cannot conceive that water is  $H_3O$ . When you imagine that water is something other than  $H_2O$ , say  $H_3O$ , you are not imagining water but something *like* water. You can merely conceive that something looking like water is  $H_3O$ , but not water itself. Similarly, my opponent can argue the following. The disjunctivist presupposes that the external object is constitutive of V appearance, as the chemical structure of  $H_2O$  is essential for water. Then, it is metaphysically impossible that the V appearance lacks physical objects. Consequently, the H that you imagine in the TE cannot be a V appearance that lacks a real object, but rather something *like* a V appearance without its real objects. From this, it follows that only the weak version (a) of the TE is legitimate, but not the strong version (b).

Now, the essential question here is whether the context in which the debate takes place makes the disjunctivist's point dialectically appropriate. This debate is somewhat reminiscent of the one between Moore (1959) and the skeptic. In the context in which the epistemic status of ordinary beliefs is at stake, Moore asserts "I know that I have hands", *ergo* "I know that there is an external world". Therefore, the skeptical hypotheses must be false. The problem is that Moore's response here simply begs the question against the skeptic, although in a normal conversational context such statements would be appropriate.

Similarly, in the context in which (let us say) the skeptic<sup>7</sup> raises the possibility of conceiving a V appearance which is "stripped" from the reality of its objects, the disjunctivist responds that this is not conceivable because one's present VP is essentially world-involving. I think there is something problematic about the disjunctivist's response here that very much resembles Moore's dialectical situation. In the context in which we try to determine whether VPs have physical objects by employing philosophical reflection (TE), the disjunctivist rejects the possibility of even conceiving of their position is false because they believe that it is actually true. But if one enters the debate by already deciding the nature of perception then why do we need to consider the thought experiment argument in the first place? The skeptic seems to be doing their job by demanding open-mindedness. All they ask us is to entertain the possibility that the V appearance can be "stripped" from its physical objects (if any). As far as the phenomenology of the experience is concerned, the scenario we imagine is possible.

Moreover, in the context where skepticism is at stake, the only non-question-begging guide to the real nature of perception is its phenomenological analysis. Since the argument from hallucination is all about inquiring the real nature and object of perception, then the TE should not, in fact, presuppose any claim about what is constitutive of perception. The upshot is this: the disjunctivist should not raise an objection to the TE by appealing to what perception really is without deciding the matter at the outset. At the very least (or best), I think there is a stand-off between the disjunctivist's and the skeptic's position here.

Let me now return to the main line of the TE argument. The sub-conclusion is that it is possible that appearances of Hs and VPs are co-referential/narrowest-type identical. Now I propose to move the argument one step forward. From “appearances of Hs and VPs are co-referential/narrowest-type identical” I propose to argue that their purported objects must be the same because “what is experience about” cannot be individuated apart from how the purported object presents itself. Here is the final formulation of the argument from hallucination based on the TE:

1. It is conceivable (and possible) that an H appearance is subjectively indistinguishable from a V appearance.
2. It is possible that H appearance exactly matches V appearance by being co-referential/narrowest-type identical with it.
3. If H appearance is co-referential/narrowest-type identical with V appearance, then as far as I can tell, in both cases I am directly aware of the same purported object.
4. When hallucinating I am directly aware of a non-physical object (sense datum). [*Conclusion of the simple version of the argument from hallucination.*]
5. ∴ When having a VP I am directly aware of sense datum. [From (3) and (4).]

Now, let me discuss how this part of the TE argument can be challenged. Consider the first possible objection which concerns the premise (3). For example, you are presented with two pictures, one with ten dots and the other one with 11 dots. They appear to be the same, although you actually see two different pictures. The disjunctivist could say that this is a counterexample that shows that sameness of appearances does not imply sameness of direct objects.

Let me try to give an answer to the counterexample. One could say that appearances of dots are different although at first you are not able to detect the difference. If you really try harder, you would be able to spot the extra dot. There is a possibility of spotting the difference between the two pictures, at least in principle. This is possible because the extra dot is reflected in the appearance, although it is undetectable for people who do not have excellent discriminative abilities. Similarly, it is possible to spot the difference between a hologram of a tiger and a real tiger. But no matter how hard you try there is no chance – in principle – of detecting a difference between the appearance of a perfect H, and a veridical one. As I have already argued, the H which you conceive in the TE is perfect because in imagining an H indistinguishable from a VP, you imagine a H appearance which is co-referential/narrowest-type identical to the veridical appearance. The additional condition of the imagined scene (“there are no physical objects”) makes no change to the appearance that you imagine, whereas the extra dot in the above example adds something. And maybe this disanalogy can potentially meet the objection. An H and a VP are *in principle* indiscriminable, namely, our epistemic situation

6

I am very thankful to Howard Robinson for raising this objection.

7

The skeptic could be anyone who does not make any commitments about the nature of perceived object. Or it could be a radical skeptic who introduces the BIV hypothesis.

with respect to their discrimination cannot be improved. Then, the dot-example is a counterexample to a more generalized version (3\*):

*If x appearance is co-referential/narrowest-type identical with y appearance, then as far as I can tell, in both cases I am directly aware of the same purported object.*

However, that does not imply that an instantiation of it (premise (3)) is thereby undermined, because there is a relevant difference between the dot-example and the hallucination case.

Finally, I would like to remark that we have good justification for the thesis that the objects of H and VP must be the same. Suppose that I did everything possible in order to spot the difference between the appearance of a perfect H and that of veridical appearance. And yet, I judge that there is no difference. Am I not justified in claiming that their direct objects are the same? I believe that premise (3) is plausible.

The strong interpretation of the TE – the co-referentiality thesis – is quite strong, and it yields the conclusion that objects of an H and a VP must be the same. However, from the weaker interpretation of the TE – the narrowest-type identity thesis – it does not necessarily follow that the two appearances (i.e. their properties) must be instantiated in the same kind of object. Namely, the disjunctivist's position is still viable. Such an argument from hallucination – I agree with Robinson – is inconclusive.

## Conclusion

The thought experiment argument was motivated by considerations that rendered the claim that subjective indistinguishability implies sameness of objects in the argument from hallucination questionable. I introduced the discussion about the common factor and disjunctivism to place the TE argument in the context of contemporary discussion on philosophy of perception. I focused on proposing a way to narrow the gap between subjective indistinguishability and the sameness of objects thesis through a TE. This way I tried to answer Putnam's sorites-based objection as well as few other problems, such as the question of whether it is really possible to imagine a perfect hallucination (disjunctivist), the chiliagon counterexample, and other related worries. There are many other respectable views and objections to the argument from hallucinations that did not get my attention. However, analyzing conditions involved in the thought experiment argument from Philosophers' hallucinations has been helpful in showing that some objections, typically raised concerning subjective indistinguishability thesis and sameness of objects, do not apply in this case.

## Literature

- Byrne A.; Logue H. (2009): "Introduction", in: A. Byrne, H. Logue (eds.), *Disjunctivism. Contemporary Readings*, pp. vii–xxix. Cambridge: MIT Press.
- Byrne A.; Logue H. (2008): "Either/Or", in: A. Haddock, F. Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*, pp. 57–94. Oxford: Oxford University Press.
- Crane. T. (2005): "The Problem of Perception", *Stanford Encyclopedia of Philosophy* (2005). Available at: <https://plato.stanford.edu/entries/perception-problem/>

(accessed on 3 November 2025).

- Farkas, K. (2006): "Indiscriminability and the sameness of appearance", *Proceedings of the Aristotelian Society* 106 (2006), pp. 207–227.
- Huemer, M. (2001): *Skepticism and the Veil of Perception*, Lanham: Rowman & Littlefield Publishers.
- Husserl, E. (1931): *Ideas. General Introduction to Pure Phenomenology*, transl. W. R. Boyce Gibson. London: George Allen & Unwin Ltd.
- Johnston, M. (2004): "The Obscure Object of Hallucination", *Philosophical Studies* 120 (2004) 1–3, pp. 113–183.
- Le Morvan, P. (2004): "Arguments against Direct Realism and How to Counter Them", *American Philosophical Quarterly* 41 (2004) 3, pp. 221–234.
- Martin, M. G. F. (2004): "Uncovering Appearances", *UCL – Mike Martin*. Available at: <http://www.homepages.ucl.ac.uk/%7Eeuctymfm/chapters.htm> (accessed on 3 November 2025).
- McDowell, J. (1982): "Criteria Defeasibility and Knowledge", *Proceedings of the British Academy* 68 (1982), pp. 455–479.
- Moore, G. E. (1959): *Philosophical Papers*. London: George Allen & Unwin Ltd.
- Neta, R. (2008): "In Defence of Disjunctivism", in: A. Haddock, F. Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*, pp. 311–329. Oxford: Oxford University Press.
- Price, H. H. (1932): *Perception*. London: Methuen & Co. Ltd.
- Putnam, H. (1999): *The Threefold Cord*. New York: Columbia University Press.
- Robinson, H. (2013): "The Failure of Disjunctivism to Deal with 'Philosophers' Hallucinations'", in: F. Macpherson, D. Platchias (eds.), *Hallucination. Philosophy and Psychology*, pp. 313–330. MIT Press.
- Robinson, H. (1994): *Perception*. London: Routledge.
- Smith, A. D. (2008): "Disjunctivism and Discriminability", in: A. Haddock, F. Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*, pp. 181–204. Oxford: Oxford University Press.
- Smith, A. D. (2002): *The Problem of Perception*. Cambridge (MA): Harvard University Press.
- Thompson, B. J. (2008): "Representationalism and the Argument from Hallucination", *Pacific Philosophical Quarterly* 89 (2008) 3, pp. 384–412, doi: <https://doi.org/10.1111/j.1468-0114.2008.00327.x>.
- Williamson, T. (2000): *Knowledge and its Limits*. Oxford: Oxford University Press.

**Ksenija Puškarić**

### **Argument iz halucinacije**

#### **Sažetak**

U ovom članku branim argument iz halucinacije temeljen na zamislivosti filozofovih halucinacija koje ne uključuju nikakvu fenomenalnu razliku u odnosu na istinsku percepciju. Riječ je, u biti, o argumentu na osnovi misaonog eksperimenta koji brani tezu da subjektivna nerazlučivost tih dvaju stanja podrazumijeva istovjetnost njihovih objekata. Razmatram protuprimjere i disjunktivističke prigovore upućene na ovu inferenciju te raspravu proširujem na uvjete koji su uključeni u zamišljanje filozofovih halucinacija. Dovoljnim sužavanjem 'jaza' između teze o subjektivnoj nerazlučivosti i teze o istovjetnosti objekata odgovaram na Putnamov prigovor temeljen na fenomenalnom soritu i na srodne probleme. Detaljnija analiza uvjeta uključenih u

*misaoni eksperiment pokazuje da neki uobičajeni prigovori na inferenciju od teze o subjektivnoj nerazlučivosti do teze o istovjetnosti objekata gube svoju uvjerljivost.*

**Ključne riječi**

percepcija, halucinacija, disjunktivizam, osjetni podaci, misaoni eksperiment, fenomenски soriti

**Ksenija Puškarić**

**Das Argument aus Halluzination**

**Zusammenfassung**

*In diesem Aufsatz verteidige ich das Argument aus Halluzination, das auf der Vorstellbarkeit philosophischer Halluzinationen fußt, die keinerlei phänomenalen Unterschied zur echten Wahrnehmung aufweisen. Im Wesentlichen handelt es sich hierbei um ein auf Gedankenexperimenten basierendes Argument, das die These verteidigt, die subjektive Ununterscheidbarkeit dieser beiden Zustände impliziere die Gleichheit ihrer Objekte. Ich setze mich mit Gegenbeispielen und disjunktivistisch geprägten Einwänden gegen diese Schlussfolgerung auseinander und erweitere die Diskussion auf die Bedingungen, unter denen philosophische Halluzinationen vorstellbar sind. Indem ich die „Lücke“ zwischen der subjektiven Ununterscheidbarkeit und der These von der Gleichheit der Objekte hinreichend verengte, entgegne ich Putnams auf dem phänomenalen Sorites beruhenden Einwand sowie weiteren damit zusammenhängenden Problemen. Eine genauere Analyse der im Gedankenexperiment einbegriffenen Bedingungen zeigt, dass einige häufig vorgebrachte Einwände gegen die Schlussfolgerung von der These der subjektiven Ununterscheidbarkeit zur These von der Gleichheit der Objekte an Plausibilität verlieren.*

**Schlüsselwörter**

Wahrnehmung, Halluzination, Disjunktivismus, Sinnesdaten, Gedankenexperiment, phänomenaler Sorites

**Ksenija Puškarić**

**L'argument tiré de l'hallucination**

**Résumé**

*Dans cet article, je défends l'argument tiré de l'hallucination, fondé sur la concevabilité des hallucinations philosophiques qui ne présentent aucune différence phénoménale avec la perception authentique. Il s'agit essentiellement d'un argument basé sur une expérience de pensée, défendant la thèse selon laquelle l'indistinguabilité subjective de ces deux états implique l'identité de leurs objets. J'aborde les contre-exemples et les objections de type disjonctiviste à cette inférence, et j'élargis la discussion aux conditions impliquées dans la conception des hallucinations philosophiques. En resserrant suffisamment « l'écart » entre l'indistinguabilité subjective et la thèse de l'identité des objets, je réponds à l'objection de Putnam fondée sur le sorites phénoménal ainsi qu'à d'autres problèmes connexes. Une analyse plus approfondie des conditions de l'expérience de pensée montre que certaines objections fréquemment avancées contre l'inférence de l'indistinguabilité subjective à l'identité des objets perdent de leur plausibilité.*

**Mots-clés**

perception, hallucination, disjonctivisme, donné sensible, expérience de pensée, sorites phénoménal