ORIGINAL RESEARCH

On the Difference Between Realistic and Fantastic Imagining

Christopher Gauker[1]

Abstract

When we imaginatively picture what might happen, we may take what we imagine to be either realistic or fantastic. A wine glass falling to the floor and shattering is realistic. A wine glass falling and morphing into a bird and flying away is fantastic. What does the distinction consist in? Two important necessary conditions are here defined. The first is a condition on the realistic representation of spatial configuration, grounded in an account of the imagistic representation of spatial configuration. The second is a condition on the manner in which realistic courses of mental imagery may be grounded in remembered perceptions. This is defined in terms of an account of the representation of comparative similarity.

1 The Problem

We commonly draw a distinction between mental imagery that is realistic and mental imagery that is fantastic. If I imagine knocking a wine glass full of wine off a shelf, and imagine it falling and shattering and splattering wine all around the point of impact, then that course of imagination will count as realistic. If the wine glass actually fell, I would not be surprised to see exactly that happen. But if I imagine a wine glass falling from a shelf and on the way down turning into a bird and flying away, then I know that what I have imagined is fantastic. If that were to happen, I would not believe my eyes.

We utilize the distinction between realistic and fantastic imagination in solving problems by means of mental imagery. If I want to wrap a box in gift wrap and need to cut a piece of wrapping paper from a roll, then I can use my imagination to determine how big the piece has to be in order to fully cover the box. If I imagine a piece of paper *so* big, then I can realistically imagine completely wrapping the box with it. But if I imagine a piece of paper only *so* big and imagine myself completely

✉ Christopher Gauker
  christopher.gauker@sbg.ac.at

1 Department of Philosophy, Faculty of Cultural and Social Sciences, University of Salzburg,
  Franziskanergasse 1, 5020 Salzburg, Austria

wrapping the box with it, then I am imagining fantastically. So I cut a piece that I can realistically imagine wrapping the box with.

Or suppose I am an employee in a restaurant. I bring a tray of clean glasses from the dishwasher to a counter, intending to transfer them to a cupboard. But I set the tray on the counter in such a way that almost half of it is not in contact with the surface. In unloading the glasses from the tray, I can take them first from the front or I can take them first from the back. If I imagine taking them from the front while the tray remains resting on the counter, then that is realistic. But if I imagine taking them from the back while the tray remains resting on the counter, then that is fantastic (because the glasses in the front will cause the tray to tip up and spill the remaining glasses onto the floor). So I take the glasses first from the front.

What I am calling imagining consists of forming mental images. I will take for granted that it is clear enough what a mental image is, although a precise definition might be hard to achieve. Roughly, a mental image is a representation that shares a format with perceptions but which differs from a perception in that perceptions are generated exogenously, in direct response to the properties of the object or scene perceived, while mental images are generated endogenously. A mental image can be conceived as a "picture in the head", but the metaphor is in many ways misleading. What I am calling a mental image need not be static, like a still photograph. It may represent a continuous sequence of events over a period of time, like a movie, in which case the mental imagery occurs over time in a sequence corresponding to the sequence of the events represented. Moreover, as we will see, a mental image may embody a kind of analysis of the represented scene in a way a picture does not. I will refer to an episode of temporally evolving mental imagery interchangeably as a *course of imagination* or as a *mental movie.* There may be auditory, tactile and other sorts of images, corresponding to sensory modalities other than vision, and a distinction between realistic and fantastic may pertain to these as well, but for present purposes I will consider exclusively visual mental imagery.

The distinction between realistic and fantastic, as I understand it, is a subjective distinction and subject to learning. What is fantastic for one thinker, may be realistic for another, who has had different experiences. If you have seen a jumbo jet sailing through the air, then a course of imagining representing such a thing will be realistic. But if I have never seen this, such a course of imagination might still be fantastic for me. Despite being subjective, the distinction is still a distinction between what a thinker ought to treat as realistic in light of the thinker's own experience and what the thinker ought to treat as fantastic in light of that experience and not merely the distinction between what the thinker actually does take to be realistic and what the thinker does not. Moreover, the distinction is not defined in terms of metacognition. What makes a course of imagination realistic is not that the thinker *judges* it to be. A judgment to the effect that a course of imagination is realistic is *true* if and only if the course of imagination has the properties that make it realistic for the judging agent.

The distinction between realistic and fantastic courses of imagination is not a distinction between propositions or between thoughts that represent things as belonging to kinds. I will take for granted that mental images do not bear propositional contents and do not classify things as belonging to kinds. As a picture of an apple

does not literally say of something *that* it is an apple, so too a mental image of an apple does not represent anything *as* an apple. It is still controversial in philosophy whether perceptions have propositional content or represent things as belonging to kinds. Those who claim that they do would presumably say that nonperceptual mental images do so as well. But it has often enough been argued, by me and others, that perceptions do not represent objects as belonging to kinds that it is fair for me to set this issue aside in this paper and take for granted that perceptions and nonperceptual mental images alike do not have propositional content.[1]

Even if imagistic representations do not have propositional content, it might be allowed that a relation of conformity between imagistic representations and propositions could be defined. So it might be thought that the distinction between realistic and fantastic courses of imagination could be defined thus: Realistic courses of imagination are those that conform to the propositional content of our antecedent beliefs; fantastic courses of imagination are those that do not conform. Call this the *belief theory* of the distinction between realistic and fantastic courses of imagination.[2] The theory has to be that realistic courses of imagination are those that conform to our *antecedent* beliefs, that is, those that the thinker possesses before constructing the course of imagination in question. For any sufficiently detailed description of that which an elaborate course of imagination represents, we might find, upon contemplating what we have imagined, that we believe that nothing that conforms to that description will ever happen. Even many realistic courses of imagination will represent courses of events that we expect will not occur just as we have imagined them. So we cannot say that a course of imagination counts as fantastic just for that reason. In particular, if we find, upon contemplating a fantastic course of imagination, that we believe that nothing like that will ever happen, then it is not the belief that makes the course of imagination fantastic.

This belief theory can be doubted on the grounds that there will not be enough beliefs to do the job. When we contemplate a course of imagination that we regard as fantastic, we may have no antecedent beliefs that rule it out, just because we had never before contemplated anything like what we imagine. Since a realistic course of imagination does not have to represent a closed system, we cannot confine realistic

---

[1] The thesis that I am objecting to can be expressed in various ways. Some, e.g., Glüer (2009), will indeed claim that perceptions have propositional content. Others, e.g., Siegel (2006), tend simply to say that perceptions represent objects as belonging to kinds. Yet others, e.g., Quilty-Dunn (2019), may say that perceptions have a discursive format. The present point is that the realistic-fantastic distinction is not a distinction between representations of any of these kinds. The critics include myself (Gauker 2011, 2012, 2017), Brewer (2006), Crane (2009), and Echeverri (2017).

[2] The belief theory in this version seems to be the position of Langland-Hassan (2016), Kind (2018) and Dorsch (2016). None of these authors proposes to define a realistic course of imagination as one that conforms to one's beliefs, but each of them distinguishes realistic, or reliable, courses of imagination from fantastic, or unreliable, courses of imagination, and in doing so appeals to conformity to beliefs and in doing so draws the distinction in no other way. Langland-Hassan eschews the term "belief", but still describes his "lateral constraints" as "stored generalizations" (2016, p. 71). Dorsch likewise avoids describing the constraints as *believed*, but he nonetheless writes of "truth-preserving constraints" that we "impose" on our acts of imagining (2016, p. 99). What Kind says is that her beliefs "infuse" her imaginings and so act as constraints (2018, p. 243; see also her "reality constraint" and "change constraint" in her 2016, p. 151).

courses of imagination to those that conform to our beliefs about what will happen in a closed system. But we do not have beliefs that comprehensively define the sorts of events that might intervene in an otherwise closed system and the effects that they might have. Such events might include a sudden gust of wind that blows a falling leaf in a new direction, a pedestrian who steps into the road, causing the driver to step on the brakes, or a lapse in the brain of a farmer that causes her to leave the gate to the goat pen open.

Although I thus deny that the realism of a course of imagination *consists* in its conformity to beliefs, I grant that a course of imagination might be fantastic due to inconsistency with one's antecedent beliefs. A person might realistically imagine that if he equips himself with bird-like wings and flaps real hard, then he will fly like a bird. But after he tries it and fails, he may form the belief that people with birds wings do not fly. After that, his mental movie of himself flying like a bird will no longer be realistic.

The distinction between what is realistic and what is fantastic in my sense does not map into any of the distinctions between possible and impossible that are current in the philosophical literature. While the distinction between realistic and fantastic is subjective, what is metaphysically or physically possible ought to be the same for everyone. Moreover, the distinction is also not a distinction between what the thinker *regards* as possible, metaphysically or physically, and what the thinker *regards* as impossible, because many courses of imagination that are fantastic for a thinker might nonetheless represent events that the thinker regards as possible, metaphysically or even physically. (An image of a horse that communicates in English is fantastic, but such a horse need not violate the laws of physics.) Further, the realistic courses of imagination are also not just those that are *epistemically* possible, that is, consistent with an epistemic background of some sort. (For one of many treatments of epistemic possibility, see Kratzer 1977). While a course of imagination counts as fantastic, I have granted, if it conflicts with one's antecedent beliefs, the converse does not hold. A course of imagination that does not conflict with any antecedent beliefs may still be fantastic.

We can now see that the distinction between realistic and fantastic courses of imagination does not concern the meaning of any modal connectives, such as "possibly". For first, the distinction is not a distinction between propositions of the kind that the operators that modal connectives express operate on. And second, the distinction does not map onto any of the distinctions between possibility and necessity that are current in the literature, whether these are distinctions of metaphysical, physical, or epistemic possibility and impossibility, not to mention deontic, bouletic, etc., varieties of possibility. This is not to deny that theorists concerned with some of these other sorts of possibility might legitimately appeal to mental imagery in explicating the distinction of interest to them (e.g., as Gregory (2019) does).

The distinction between realistic and fantastic is also not the distinction between probable (or not improbable) and improbable. For any rich course of perceptual experience there will be a description of what is perceived such that the probability of something's meeting that description is very low. For example, what I perceive when driving down a busy street may be described in such detail that the probability that an event so described would happen (prior to its happening) is very low. But

that will not mean that we take ourselves to be hallucinating rather than perceiving. Likewise, for any sufficiently rich course of imagination, there is a description of it such that it is highly improbable that anything that fulfills that description will ever happen. Still, it might qualify as realistic. There may be cases in which we antecedently believe that a given kind of event will not happen, just because it is highly improbable, such as flipping a coin one hundred times and getting heads each time. In those cases, a course of imagination representing such an event might also be fantastic due to a conflict with antecedent beliefs. But in other cases, a given course of imagination might, under some description, be improbable and nonetheless qualify as realistic. For example, if I imagine driving down a busy street, then there will be a description of what I imagine such that, so described, what I imagine is improbable; but that course of imagination might nonetheless count as realistic.

Despite the remaining unclarity in the distinction, I will assume that the examples make it meaningful to ask the following question: *What does the distinction between realistic and fantastic imagination consist in?* My objective in this paper is to produce a definite, albeit rather abstract answer. The answer will rest on a general account of imagistic representation; so I will begin with that. I will then define a necessary condition on realistic courses of imagination regarding spatial configuration and two procedures for constructing mental movies on the basis of remembered perceptions. Finally, I will identify the realistic courses of imagination with those that can be produced by means of those procedures and which meet a to-be-defined condition concerning the representation of spatial configuration and do not contradict our antecedent beliefs.

2 Imagistic Representation

Mental representations are states of, or events in, the brain. So in principle we might define the distinction between realistic and fantastic courses of imagination wholly in neurophysiological terms. In practice, however, we can refer to mental representations only in terms of what they represent. Moreover, we might want a theory that generalizes across species and even across life-forms and robot-forms. A theory as general as that would have to be formulated in terms of what is represented, not in terms of intrinsic states of the thinking machine. For both of these reasons, we have to begin with an account of the way in which mental imagery represents.

2.1 The Representation of Configuration

An imagistic representation of a scene parses the scene into individual objects, parses those objects into such things as surfaces, edges, and perhaps major axes, and represents the spatial relations between these entities. These are not three separable processes. The representation of the individual objects goes hand-in-hand with the representation of the individual surfaces and edges. The representation of surfaces and edges goes hand-in-hand with the representation of the spatial relations between them. Inasmuch as perceptual processing is the product of learning and

a perceptual representation may result from sensory inputs over a period of time, the relations represented might also include relations in depth relative to the line of sight, so that even partially occluded parts of an object are so represented. The challenge in explaining the imagistic representation of configuration is to identify the relation between the elements of the representation and the elements of the scene represented by virtue of which the former counts as a representation of the latter.

Let us say that a perceptual representation of configuration consists of a construction of *mental markers* in *relations* to one another. These markers are elements of the perceptual representation itself (having some kind of neurological nature). Intuitively, we can think of these markers as representing the basic elements of the scene. There may be one for each object and one for each surface of each object. Certain relations between the markers will represent relations between the basic elements of the scene. (I will treat properties as 1-place relations.) For instance, one of these relations R might be such that if one marker x stands in relation R to marker y, then the object that x represents is represented as being *to the right of* the object that y represents, so that the representation consisting of x in relation R to y is accurate only if the object that x represents is to the right of the object that y represents. For another example, a relation J may be such that if a marker u stands in relation J to a marker v, then the surface that u represents is represented as joined along one edge at a 45° angle to the surface that v represents. Exactly how the image should be analyzed into elements and relations between elements, and exactly what entities and relations need to be represented in a perceptual representation is a question that I cannot take up here.[3]

More generally, then, our account of the representation of configuration will involve two functions, a function Π, which takes us from n-ary relations between mental markers into n-ary relations between external objects or features, and a function h, which takes us from mental markers to external particulars (e.g., objects, surfaces and edges). (See Fig. 1.) Π is analogous to a projection according to scale, which tells us how much distance on a paper map corresponds to a given distance on the terrain mapped. In the present instance, Π cannot be literally a projection according to scale because the representation in question is not a paper map or literally a picture of any kind but a state of the brain. In the example in the previous paragraph, $\Pi(R) = being\ to\ the\ right\ of$ and $\Pi(J) = being\ joined\ at\ an\ edge\ at\ a\ 45°\ angle$. Continuing the map analogy, h is like the function that takes us from spots on the map to places in the terrain mapped (e.g., a town) and from regions on the map

---

[3] Computer models of vision, such as that in Kubilius et al. (2014), aim to explain how such parsing of a visual representation of a scene might be possible. There are many studies, e.g., Singh and Hoffman (2001), seeking to identify various aspects of the way that humans actually do it. For a review of studies suggesting that the three-dimensional structure of perceived scenes is represented at some level of processing, see Hummel (2013). For evidence that configuration is already represented at the level of perception, see Green (2019). Although there is ongoing work of this kind attempting to explicate the nature of visual parsing, I am not aware of any attempt since Palmer (1978) to explain in a general way the representation relation between an imagistic representation and an external scene. (Palmer does not explicitly confine himself to imagistic representations. He claims to extend his account to "propositional representations", pp. 294–295, but he does not address the semantics of logically complex sentences.)
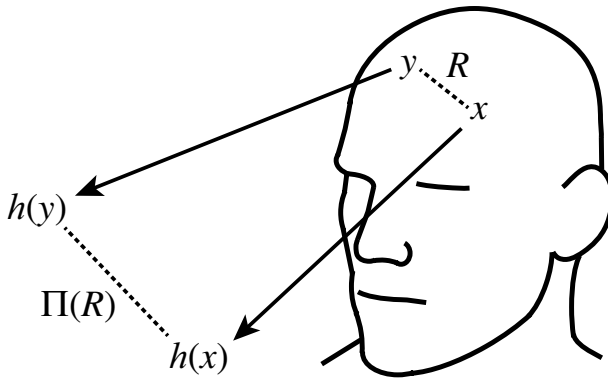
Fig. 1 Representing configuration. x and y are mental markers, R is a relation between them, $h(x)$ and $h(y)$ are parts of an external configuration, which h maps x and y into, and $\Pi(R)$ is a relation between those parts, which Π maps R into

into regions on the terrain (e.g., a lake). The account of perceptual representation of configuration will take the form of an account of *accurate* representation. The account of accurate representation will proceed in two stages. First, we will define accuracy *relative* to arbitrarily chosen functions Π (from relations into relations) and h (from mental markers into external entities). The second step would be to identify the functions Π^* and h^* in terms of which accuracy *simpliciter* should be defined.

An imagistic representation, consisting of mental markers $x_1, x_2, \ldots, x_n$, is a *completely accurate* representation of the configuration of a scene relative to Π and h if and only if h is a mapping of $x_1, x_2, \ldots, x_n$ into elements of the scene and for all n-ary relations M in the domain of Π, $x_1, x_2, \ldots, x_n$ (in that order) stand in M if and only if $h(x_1), h(x_2), \ldots, h(x_n)$ (in that order) stand in $\Pi(M)$. In general, h is a *homomorphism* relative to Π if and only if for all n, for all n-ary relations M in the domain of Π, for all $x_1, x_2, \ldots, x_n$ in the domain of h, $x_1, x_2, \ldots, x_n$ (in that order) stand in M if and only if $h(x_1), h(x_2), \ldots, h(x_n)$ (in that order) stand in the n-ary relation $\Pi(M)$. (h will be an *isomorphism* relative to Π if and only if it is a homomorphism relative to Π and also a one-to-one function.) So we can reformulate the definition of accuracy in terms of homomorphisms thus: An imagistic representation consisting of mental markers $x_1, x_2, \ldots, x_n$ is a *completely accurate* representation of the configuration of a scene relative to Π and h if and only if h is a mapping of $x_1, x_2, \ldots, x_n$ into elements of the scene that is a homomorphism with respect to Π *in the region* (of the domain of h) consisting of $x_1, x_2, \ldots, x_n$. For example, if a and b are two mental markers constituting a completely accurate representation relative to Π and h, then if a stands in relation R to b (from the example above), then $h(a)$ will be an external object to the right of the external object $h(b)$. Obviously, a representation can be completely accurate, as far as it goes, without completely representing every aspect of the thing represented.

Complete accuracy, as here defined, is a limiting case that is not often realized. Departures from the ideal can be countenanced by countenancing the fact that an actual relation between the elements of a scene may only approximate to the relation

into which Π maps the relation between the mental markers that h maps into elements of the scene. For instance, a representation that represents three points as lying on a straight line (which means that Π maps the relation between the markers that h maps into the points into the relation of lying on a straight line) may be nearly accurate if the line connecting the three points represented (into which h maps the markers composing the representation) is in fact nearly straight.

So far we have only defined accuracy relative to arbitrarily chosen functions Π and h. To define accuracy simpliciter, we need to discharge the relativity to Π and h by identifying the functions Π^* and h^* such that accuracy relative to Π^* and h^* is accuracy *simpliciter*. Here I will only make two observations that might be put to use in a full answer. The first of these is that we can think of the mapping h^* of mental markers into external particulars as describing a particular perceptual mechanism. So conceived, the output of the function h^*, given a mental marker x as input, is the external entity the appearing of which before the senses generates, in accordance with the mechanism, the mental marker. (So the cause of the marker is the thing that h maps the marker into.) I say that it is the *appearing* of the external entity before the senses that generates the marker, because I assume that a cause has to be an event. Appearing in this sense is an event in which something that was not present to the senses becomes present to the senses. A paradigm would be the mechanism that begins with light being reflecting from an edge between a light-colored region and a dark-colored region into the animal's eyes and that ends with the firing of a neuronal edge-detector (compare the classic edge detectors of Hubel and Wiesel 1962).

The second observation is that the relations between mental particulars play a role in determining the effect that the possession of these mental markers will have on the animal's behavior. For instance, it might be the case that if mental marker x stands in (mental) relation R to mental marker y, then $h^*(x)$ will be treated in action as if it were to the right of $h^*(y)$. Moreover, treating external entities as standing in certain relations may have an effect on the animal's success in meeting its needs. Thus, if the animal *needs* object $h^*(x)$ more than it *needs* object $h^*(y)$, then, as a consequence of x's standing in relation R to y, the animal will reach to the right and not to the left and thereby succeed in meeting its needs. More generally, if mental markers $x_1, x_2, \ldots, x_n$ (in that order) stand in n-ary relation M to one another, and Π^* and h^* are the functions in terms of which accuracy simpliciter is defined for this kind of animal, then $h^*(x_1), h^*(x_2), \ldots, h^*(x_n)$ (in that order) will be treated in action as if they stood in relation $\Pi^*(M)$, and treating them so may have, on the whole, a positive effect in enabling the animal to meet its needs.

The task of identifying Π^* and h^*, the functions such that accuracy relative to them is accuracy simpliciter, will make use of the above two observations. In view of these two observations it should be possible to identify Π^* and h^*, for a given kind of animal, with *the* functions Π and h *such that* the fact that it is biologically normal for creatures of that kind to represent things accurately relative to Π and h accounts for the fact that animals of that kind are able to meet their biological needs. How exactly to complete this strategy, however, is beyond the scope of this paper. We will want to ensure that the answer identifies Π^* and h^* uniquely, and we will want to ensure that the answer allows that an animal is capable of representing inaccurately as well as representing accurately.

Endogenously generated mental markers do not represent any actual scene, or do so at most accidentally. Nonetheless, we may assume that the mental markers and the relations between them that constitute endogenously generated mental images are, at some level of neurological properties and relations, the same in kind as those that make up exogenously generated perceptual representations. In the case of a non-perceptual mental image and a mental marker x that is part of it, $h^*(x)$ does not really exist. But though a fictitious object cannot literally be denoted and cannot, in an extensional sense, even be described (since there *is* nothing so described), we can nonetheless produce *descriptions* of fictitious objects that we can use to produce a description of a fictitious scene (in an intensional sense of "description of"). In particular, we theorists of perception can use a description of the mechanism that h^* corresponds to in order to produce a description of a nonexistent scene. Where $x_1, x_2,\ldots, x_n$ are the mental markers that make up a given, endogenously-generated mental image and in the image they stand in relation M, the nonexistent representata of a mental image can be described thus: a scene in which $h^*(x_1), h^*(x_2),\ldots, h^*(x_n)$ (in that order) stand in relation $\Pi^*(M)$. Though $h^*(x_1), h^*(x_2),\ldots, h^*(x_n)$ do not actually exist, the expressions "$h^*(x_1)$", "$h^*(x_2)$",…, "$h^*(x_n)$", considered as describing the input to a mechanism, are still meaningful singular terms, albeit nondenoting. Of course, we might have more commonplace descriptions for the kinds of things meeting these descriptions, such as "cube next to a sphere" or even "flying turtle". In light of this, I will consider myself free in what follows to refer to *the configuration of objects and surfaces that an imagistic representation represents* even when that representation is endogenously generated and not a perception of any actual scene.

2.2 The Representation of Similarity

Our perceptions and images do not merely represent the configurations of surfaces, they also represent colors, textures, temperatures and other qualities that less readily come to mind, such as jerkiness of motion. This does not mean that they attribute properties or kinds to individuals. What it means is that they measure the locations of things along various dimensions. It is not very plausible that the brain produces numerical measures of the location of an object along various dimensions of variability. But it is quite plausible that the brain makes comparisons among things along various dimensions. For instance, object x might be rated as more like y than like z with respect to color or with respect to jerkiness of motion. In just that sense we can say that the brain "measures" the location of an object along a given dimension.

Accordingly, I will suppose that the second aspect of imagistic representation is representation of similarities. The representation takes the form of a point, or, as I will say, a *mark* in a *perceptual similarity space*. What the mark represents is the location of an object or an arrangement of objects in a many-dimensional *objective quality space*.[4] The location of the mark in perceptual similarity space depends on

---

[4] The concept of a similarity space has been used prominently in philosophical writings in order to explicate the nature of concepts (Churchland 1989; Gärdenfors 2000). So I should say that I do not foresee any such use for it. For my critique of such applications, see Gauker (2007, 2011).

how the mark is produced by perceptual systems. The location of the object in objective quality space is determined by where the object or arrangement actually lies on the various dimensions. For simplicity, we may think of the objective quality space as containing only those dimensions that correspond to dimensions of perceptual similarity space. Which dimensions perceptual similarity space actually contains is an empirical question. There might be hundreds or even thousands of them. Not every mark in perceptual similarity space will specify a location on every dimension of perceptual similarity space; a mark will be only a partial vector. That the perceptual systems place mark x closer to mark y than to mark z may be interpreted as the systems' representing x as more like y than like z (collectively with respect to the dimensions measured). Although we should not assume that the brain produces a numerical measure of distance along each of the dimensions of perceptual similarity space, we theorists may find it useful to introduce numerical measures that we can use to represent the relative distances between representations.

A mark in perceptual similarity space is the product of a great deal of processing. We should not suppose that the comparative location of a thing along each of the dimensions of perceptual similarity space is immediately read off of the stimulation of the retina or the inner ear. A mark is itself a representation of a configuration, of the kind defined in the previous section. So, in a sense, a *mark* is itself structured, composed of what I earlier called *markers*. Accordingly, the representation of location in objective quality space may take account of occluded portions of objects. Moreover the processing that produces a mark may integrate inputs that take place over time or which come from several sensory modalities. Accordingly, the processing may exhibit perceptual constancies. For instance, a perception of a single object may, within limits, produce the same mark on the various color dimensions despite varying lighting conditions. In addition, we can allow that there is room for learning in the production of marks in perceptual similarity space.

Although representation of configuration is, on the present account, an independent aspect of perceptual representation, configurations can also be compared for similarity as such. A pair of cylinders joined by a hinge can be compared to other such pairs with respect to the angle of the two cylinders at the joint and with respect to the length and the breadth of the component cylinders, so that we may say of one such configuration x that it is more like another such configuration y than it is like a third such configuration z. Of course, one configuration, e.g., a configuration of furniture in a dining room, may be more or less incomparable to another, e.g. the body parts of a squirrel. So we should suppose that perceptual similarity space contains a number of sub-spaces such that marks in distinct subspaces may be more or less incomparable on the dimensions that form distinct subspaces. The subspace in which dining room sets are represented will be distinct from the subspace in which the postures of four-footed animals are represented.

As in the previous section, my account of representation will take the form of an account of accurate representation. To this end, let us distinguish between three points, one in perceptual similarity space and two in objective quality space. The first of these is the location of a mark in perceptual similarity space. To this point there corresponds two points in objective quality space. One of these is the point where the object or arrangement of objects that qualifies as *the cause* of the mark
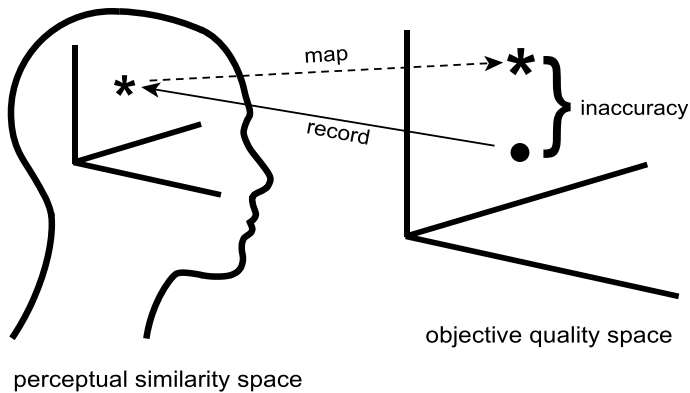
Fig. 2 Inaccurate representation in perceptual similarity space

actually lies in objective quality space. The other point in objective quality space is the point into which the point in perceptual similarity space occupied by the mark is *mapped*, in a sense to be explained presently. In terms of these two points in objective quality space we can define the accuracy of the mark in perceptual similarity space, thus: A mark in perceptual similarity space is *accurate* to the extent that the point in objective quality space that the mark is mapped into is *near* to the point in objective quality space that the cause of the mark occupies. In other words, the accuracy of a mark is inversely proportional to the distance between the point occupied by the cause of the mark and the point that the mark is mapped into. (See Fig. 2.)

In the previous paragraph I spoke of *the cause* of a mark in perceptual similarity space. Here is how we can select the relevant cause from the whole sequence of events leading up to the recording of a mark in perceptual similarity space. As I explained in the previous section, perceptions represent the configurations of certain sorts of things by virtue of the functions Π^* and h^*. Call these configured things *navigables*. The link in the chain of causes and effects leading up to the given mark in perceptual similarity space that counts as *the cause* of the mark in my sense is the nearest link to the mark that is the appearing of such a navigable. As before, I say that the cause is the *appearing* of the navigable, not the navigable itself, because a cause has to be an event. Nonetheless, I will simplify by speaking of an object or arrangement of objects as itself the cause of a mark.

I have distinguished between the point in objective quality space that is occupied by the cause of a mark in perceptual similarity space and the point in objective quality space that a mark in perceptual similarity space is *mapped into*. The next step would be to explain this mapping. The basic idea is that there is manner in which marks in perceptual similarity space are recorded when the perceptual systems are functioning properly. In the case of biological creatures this proper function might be defined in terms of biological norms. Then we can say that a mark in perceptual similarity space is mapped into that location in objective quality space where the object that causes the mark *would have been* if the mark had been recorded in accordance with the manner in which marks are recorded when perceptual systems

are functioning properly. I think it is fair to assume that there is a distinction between perceptual systems functioning properly and perceptual systems malfunctioning. The distinction calls for a more careful definition, but providing such a definition is beyond the scope of this paper.[5]

3 Realistic Versus Fantastic

In terms of these accounts of two aspects of imagistic representation we can draw the distinction between realistic and fantastic courses of imagination. We may distinguish between two aspects of the problem, the realism of the representation of spatial configuration and the realism of the representation of transformations.

3.1 Realistic Configuration

Not everything that we can visually imagine represents a configuration of a kind that we could accurately perceive. There are impossible figures that we can imagine but which we do not regard as realistic representations of configuration. There is a lithograph by M. C. Escher, for example, that depicts a staircase that circles back to itself ("Klimmen en dalen" 1960). A mental image depicting such a staircase should normally count as fantastic. The question that concerns us in this section is, what is the distinction between those courses of imagination that the mind takes to be realistic representations of spatial configuration and those that it does not take to be realistic representations of spatial configuration?

As we have seen in Sect. 2.1, a perceptual representation of configuration is a construction consisting of mental markers in relations to one another, and the accuracy of such a construction can be defined in terms of the functions $h*$ and $\Pi*$. So we can conceive of a *way* of constructing a perceptual representation that the mind would engage in when its representations of spatial configuration were accurate. Call this way the *accurate* way of constructing perceptual representations of spatial configuration.

This accurate way of constructing perceptual representations of spatial configuration is an ideal that can be approached but may never be reached. To the extent that the accurate way is approximated, that approximation may have to be acquired through maturation and learning in the course of interacting with the environment. At a certain point in its development, the mind might find nothing unsettling in the endlessly descending staircase in Escher's etching. Even as adults, we are subject to some systematic inaccuracies in perception (for instance, in underestimating distances from ourselves). Still, we can say that the manner of constructing perceptual representations of spatial configuration that the mind settles for is that which the mind *takes* to be the accurate way of constructing perceptual representations of
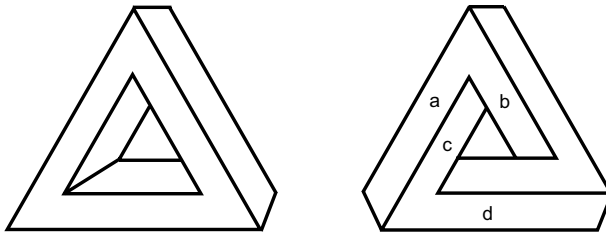

Fig. 3 A normal triangle (left) and the Penrose triangle (right)

spatial configuration. (So, as here defined, *taking* a construction to be accurate is not a matter of forming a metacognitive attitude of some kind toward it.)

We can then say that an endogenously generated mental image is a *realistic* representation of spatial configuration for a given mind if and only if it can be constructed in the same way that the mind, in this sense, *takes* to be the accurate way of representing spatial configuration in the case that the representation is a perception (except, of course, that it is not a perception). In the ideal case in which the mind has learned to perceive spatial configuration accurately, then, a mental image will be realistic if and only if it can be constructed from mental markers and relations between them in the way that, in the case of a perception, would produce an accurate representation of a spatial configuration. However, the distinction between realistic and fantastic representations of configuration is a subjective distinction inasmuch as the way of forming perceptual representations that the mind *takes* to be accurate may not in fact be a completely accurate way of forming perceptual representations.

Instead of the Escher staircase, let us take as an example of an impossible figure the Penrose triangle (see Fig. 3). The example to be considered here is not a drawing of the Penrose triangle on paper but a mental image of the Penrose triangle, which, again, is not literally a picture in the brain. Considered as an image of a three-dimensional object, and not merely as an image of lines on paper, the mental image is a fantastic representation of configuration. If we suppose —hypothesize, contrary to fact — that a mental image of the Penrose triangle is an accurate perception of an actual object in space, then it must meet the following condition: h^* maps each element of the image into an element of the object represented and Π^* maps each relation between elements of the image into a relation between elements of the object represented in such a way that h^* is a homomorphism relative to Π^* in the region of the domain of h^* containing the elements of the image. Since no such mapping of the mental markers and their relations that makes up the representation meets this condition, the representation cannot in fact be an accurate representation of an actual object in space.

That no actual object satisfies this condition may be demonstrated roughly as follows: We may suppose that Π^* maps the relation of being joined with no boundary, which holds between regions *a* and *b* (see Fig. 3, right-hand side) and between *c* and *d*, into the relation of being co-planar. (Here we let relations between elements of the drawing on paper stand in for relations between elements of the mental image, the nature of which, in our ignorance of the relevant neurophysiology, is unknown

to us.) We may suppose that Π* maps the relation of being joined with a boundary, which holds between regions *a* and *c*, into the relation of being joined at an angle. But if *a* and *b* are mapped into co-planar regions and *c* and *d* are mapped into co-planar regions, then *a* and *c* cannot be mapped into regions joined at angle. The complete proof would take into account all regions of the diagram mapped into visible regions of the hypothetical figure in space, together with general facts about the geometry of solid figures in space.

What happens when the mind deems a visual image of the Penrose triangle to be fantastic is that it executes subconsciously a version of the proof just alluded to and consequently fails to place the parts of the representation into relations to one another in the way it takes to be accurate in perceptions of spatial configurations. More generally, one case in which the mind deems an image to be a fantastic representation of spatial configuration is that in which, on the basis of an analysis of a limited portion of the image, the mind places elements $x_i, \ldots, x_j$ into relation *M* in the domain of Π*, and on the basis of an analysis of another limited portion of the image, places $x_k, \ldots, x_l$ into relation *N* in the domain of Π*, and so on, and then determines that some subset of these determinations (such as the determination that $M(x_i, \ldots, x_j)$) is inconsistent with some other one of them (such as the determination that $N(x_k, \ldots, x_l)$). That a placement of mental markers in relations to one another *really* is inconsistent means that the things that *h** maps the elements of the image into cannot, in reality, stand in the relations that Π* maps the selected relations between the elements of the image into. The mind *determines* some set of determinations to be inconsistent with some other set when it finds no way to combine them in the way they are combined in when the mind takes itself to have formed an accurate perception.[6]

3.2 Realism Through Permissible Transformations

Every realistic course of imagination must be realistic with respect to spatial configuration at each temporal stage. In this section, I will assume that all courses of imagination at issue pass this test. Within the class of courses of imagination that pass this test, only those courses of imagination will be regarded as realistic that can be constructed by two basic processes. One is the uniform translation of a perception of an event across the dimensions of perceptual similarity space. The other is a kind of linking of sequences of mental images that share endpoints.

For an example of the first process, suppose I have observed a ballet dancer dance across the stage wearing a red costume. Having seen that, I can realistically imagine her doing the same dance wearing a green costume. Or I can imagine her somewhat taller or moving somewhat faster. For an example of the second

---

[6] Readers may have seen videos of constructions by Kokichi Sugihara, in which three-dimensional objects look, from a certain perspective, like the impossible figures in Escher's drawings (and objects move along apparently impossible trajectories). If we were actually perceiving such a construction (and not watching a video), we might describe our perception as a fantastic perception, because our mind fails to construct a perception that we take to be an accurate perception of an actual scene.
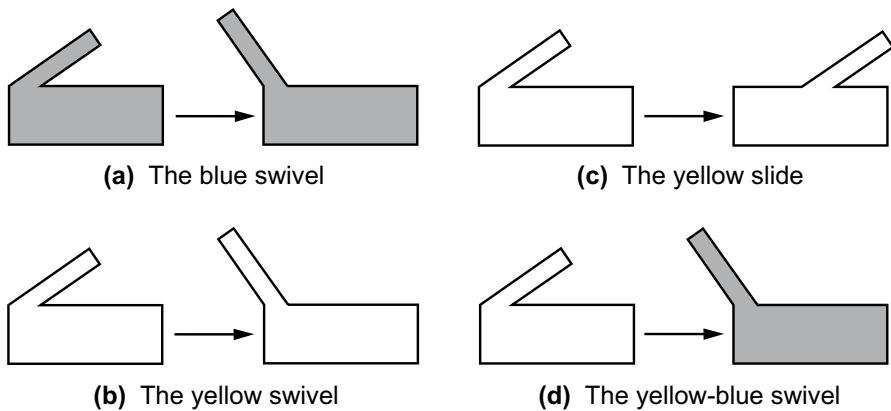
(a) The blue swivel

(c) The yellow slide

(b) The yellow swivel

(d) The yellow-blue swivel

Fig. 4 Various transformations

process, suppose I observe the ballet dancer first performing a pirouette from fifth position to fifth position, without moving the lifted leg from front to back, and then, second, performing a retiré passe, again without changing the lifted leg from front to back. Since the position from which she started the first step is the same as the position in which she ended the second step, I can crop in imagination the retiré passe that followed from the pirouette with which she began and paste it ahead of the pirouette. To formulate this thesis in a general way, I will first define *small permissible transforms* and then in terms of those I will define *permissible transforms (tout court)*.

The small permissible transforms are courses of imagination that are directly modeled on sequences of perceptions preserved in memory. Roughly speaking, a small permissible transform is a course of imagination that is *similar* to a remembered sequence of perceptions representing events that we have actually observed. But this is not quite an adequate description, because not every kind of similarity counts, as I will now explain.

Imagine a box with a blue handle attached to it, which can be swiveled up and down. Suppose we have observed the handle being swiveled over the top of the box. Call this sequence of perceptions *the blue swivel* (Fig. 4a). What we would like to say is that on the basis of the blue swivel we can form a realistic sequence of images representing the same motion but in which the box and the handle are yellow. Call this the *yellow swivel* (Fig. 4b). By contrast, since we have never observed the handle of the blue box being shifted across the top of the box, a sequence of images representing the handle of a yellow box being shifted across the top, the *yellow slide* (Fig. 4c), remains fantastic. But there is a problem. Imagine the box starts out yellow but, after the handle has been swiveled across the top, ends up blue. A sequence of images representing this transformation, call it the *yellow-blue swivel* (Fig. 4d), is in some ways even more like the original blue swivel than the yellow swivel is, because it represents the same blue color at the end as the original blue swivel did.

Nonetheless, we would like to say that the yellow-blue swivel is fantastic, not realistic, and so should not qualify as a small permissible transform.

The problem with the yellow-blue swivel is that in modeling it on the blue swivel, we have not *uniformly* transformed the model in order to create the new sequence. For that reason the yellow-blue swivel should not count as a small permissible transform. To formulate in a general way the relation that a sequence of mental images has to bear to a sequence of perceptions in order to count as a small permissible transform based on the model, we need to bring to bear our account of the representation of location in objective quality space. Roughly, we will say that a small permissible transform is the product of *uniformly translating* a sequence of perceptions across one or more dimensions of perceptual similarity space.

An observation of a sequence of events may be characterized as a sequence of marks in perceptual similarity space. Call such a sequence of marks representing a sequence of perceptions a *perceptual trail*. If we draw the smoothest possible curve through these points, that curve will be what I call a *perceptual trajectory*. Similarly, a sequence of endogenously generated marks in perceptual similarity space may constitute a mental movie representing an imaginary course of events. The smoothest possible curve drawn through such a sequence of endogenously generated marks is an *imaginary trajectory*. I will assume that many such perceptual trajectories are preserved intact and without modification in memory for use in constructing mental movies later, although I acknowledge that this may be an idealization.

Two trajectories through perceptual similarity space may *run parallel* to one another in the sense that second results from the first by, so to speak, dragging it some distance across similarity space without bending it. More precisely, two trajectories A and B through perceptual similarity space *run parallel* if and only for each of a number of dimensions there is a number such that B results from adding that number to each point on A on that dimension. More precisely, if the dimensions on which points on trajectory A are evaluated are labeled 1, 2,..., n, and each point on trajectory A is a vector $\langle a_1, a_2,..., a_n \rangle$, with a_i being the value of A on dimension i ($1 \leq i \leq n$), then B *runs parallel* to A if and only if there are real numbers $m_1, m_2,...,$ m_n (positive, negative or zero) such that $B = \{\langle b_1, b_2,..., b_n \rangle |$ for some point $\langle a_1,$ $a_2,..., a_n \rangle$ on A, $b_1 = a_1 + m_1, b_2 = a_2 + m_2,..., b_n = a_n + m_n \}$. (In the case of dimensions j along which there is no translation, $m_j = 0$.) I acknowledge that I am idealizing to the extent of supposing that for each dimension of perceptual similarity space, there is a metric that measures distance along that dimension. We have to define parallelism in terms of the trajectory that runs through the trail and not in terms of the trail itself, because we cannot assume that the number of perceptions in the perceptual trail equals the number of mental images in the imaginary trail or that the mental images are spaced along their trajectory at the same interval as the perceptions are spaced along their trajectory.

Now we can say that a sequence of endogenously generated marks in perceptual similarity space (mental images) is a *small permissible transform* of a sequence of exogenously generated marks in perceptual similarity space (a sequence of perceptions) if and only if there is an imaginary trajectory running through the former and a perceptual trajectory running through the latter such that the imaginary trajectory *runs parallel* to the perceptual trajectory (see Fig. 5). For short, a small permissible
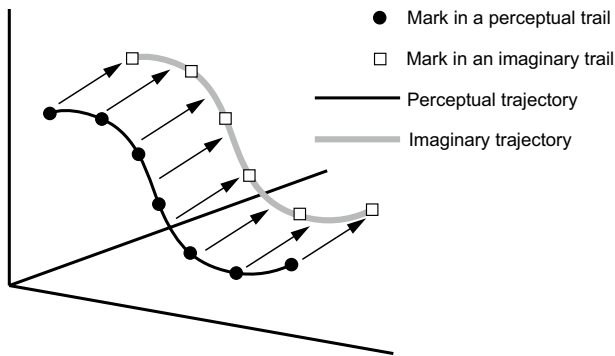
Fig. 5 A small permissible transform

transform is the product of a *uniform translation* of a sequence of perceptions across perceptual similarity space. The reason why the yellow-blue swivel is not a small permissible transform is that, while it is modeled on the blue swivel, which is a sequence of perceptions, the mental image at the start and the mental image at the end do not result from equal displacements along the color dimension(s) of the perception at the start of the blue swivel and the perception at the end of the blue swivel.

To what I just said I need to add one qualification. The values that a small permissible transform adds to the values along a given dimension of a perceptual trail must not be *too large*. Though we may have observed a man winning the 100 m dash in the Olympics, no permissible transformation will represent a man running twice as fast as that. This fuzzy restriction will mean that the boundary to be drawn between realistic and fantastic will be somewhat fuzzy. It is not obvious how to define how much is too much, and I will not explore this topic any further here. Suffice it to say that the limits will be based on past experience somehow.

Now we are in a position to define permissible transforms in general. Let us say that two small permissible transforms are *linked* if and only if the last mark in the first is the first mark in the second. Likewise, a *sequence* of small permissible transforms is *linked* if and only if for each ith member of the sequence, other than the last, the last mark in the ith permissible transform is the first mark in the $(i+1)$th member of the sequence. A mental movie counts as a *permissible transform* if and only if it consists of a sequence of linked small permissible transforms (see Fig. 6). There is here no assumption that the perceptions that serve as the models for the several permissible transforms that compose a realistic mental movie are themselves linked. The several permissible transforms that make up a realistic mental movie may be modeled on very different perceptions. For example, if I realistically imagine a dance by a child, part of it might be modeled on my observations of a dance by a grown woman, and part of it might be modeled on my observations of a dance by a grown man.

A mental movie will fail to be a permissible transform if it contains a segment that is not based on any observations. That is to say, there is no way to
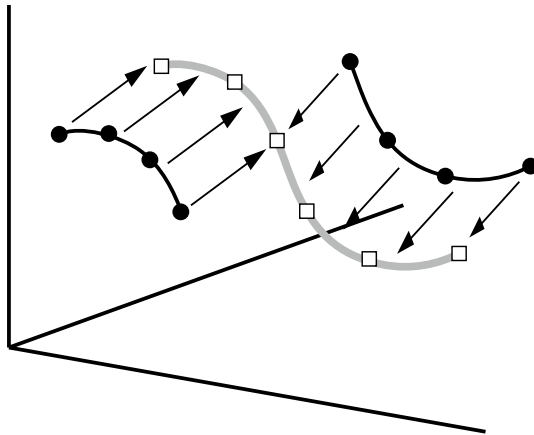
Fig. 6 A permissible transform

divide the mental movie into linked segments such that each segment is a permissible transform of a series of perceptions. For example, in my mental movie of the falling wine glass that turns into a bird, the portion of the mental movie in which the wine glass morphs into a bird is not a uniform translation across similarity space of anything that I have actually observed.

My hypothesis is that every realistic course of imagination (mental movie) is a permissible transform in the sense here defined. On this account of realistic courses of imagination, the realism of a course of imagination has to be proved, by constructing it from a linked series of small permissible transforms, the permissibility of each of which has to be established by translating a remembered perception across some dimensions of perceptual similarity space. Mental movies that cannot be demonstrated to be realistic in this way will be discounted as fantastic. Of course, a mental movie that is once regarded as fantastic may come to be regarded as realistic if observations are made that allow the mental movie to be constructed in the requisite manner.

This condition on realistic mental movies is in one respect very restrictive. It does not admit simultaneous transformations in two regions of a mental image if they are not uniform translations of transformations that have been simultaneously observed. For instance, I have seen a dog catch a tennis ball in its mouth, and I have seen a dog jumping through a hoop. But unless I have seen a dog (or some other animal) catching a tennis ball (or something like it) in its mouth while jumping through a hoop (or something like a hoop), a mental movie representing such an event will not qualify as realistic. This is a desirable consequence, because we do not know a priori what combinations of events are possible. I have seen a dancer lift his left leg without falling. And I have seen a dancer lift his right leg without falling. But I have never seen a dancer lift his left leg and his right leg simultaneously without falling. So if I imagine his lifting both his left leg and his right leg at the same time without falling, then that course of imagination should not count as realistic.

3.3 Putting it All Together

There is one last condition that a realistic course of imagining must meet: As noted in Sect. 1, a course of imagination may count as realistic only if it does not conflict with one's antecedent beliefs. Here then is my total hypothesis: A course of imagination C is realistic for a thinker P if and only if (1) C is at each moment realistic with respect to spatial configuration (Sect. 3.1), (2) C is a permissible transform for P (Sect. 3.2), and (3) the events represented in C do not contradict any of P's antecedent beliefs.

4 Problem-Solving Revisited

I said at the start that the distinction between realistic and fantastic courses of imagination is utilized in problem-solving. In my gift-wrap example, the workable imagined solution (cutting off a piece of paper *so* big) should count as realistic for me by virtue of a permissible transformation of past experiences of cutting off sheets of things and wrapping things up. In my glasses-on-the-tray example, the workable imagined solution (taking the glasses first from the front) should count as realistic by virtue of a permissible transformation of past experiences of things being removed from a surface balanced on a pivot.

The distinction between realistic and fantastic courses of imagination will be at most one of the key elements in an account of how, by means of forming mental images, we solve problems. Another part of the problem, not addressed here, is how the mind *finds* a useful realistic course of imagination. This in turn divides into two problems, that of identifying the past experiences that may serve as models and that of undertaking the useful permissible transformations. Yet another part of the problem, also not addressed here, is how the mind then puts a realistic course of imagination to work. Not every realistic course of imagination consists in imagining something the thinker can do, and even when a course of imagination represents something the thinker can do, putting that course of imagination into action may involve a process of accommodation between what the thinker imagines and what happens when he or she undertakes to act as he or she has imagined doing.

References

Brewer, B. (2006). Perception and content. *European Journal of Philosophy, 14,* 165–181.

Churchland, P. M. (1989). On the nature of theories: A neurocomputational perspective. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science 14: The nature of theories* (pp. 59–101). Minneapolis: University of Minnesota Press.

Crane, T. (2009). Is perception a propositional attitude? *The Philosophical Quarterly, 59,* 452–469.

Dorsch, F. (2016). Knowledge by imagination: How imaginative experience can ground factual knowledge. *Teorema, 35,* 87–116.

Echeverri, S. (2017). Visual reference and iconic content. *Philosophy of Science, 84,* 761–781.

Gärdenfors, P. (2000). *Conceptual spaces*. Cambridge, MA: MIT Press.

Gauker, C. (2007). A critique of the similarity space theory of concepts. *Mind and Language, 22,* 317–345.

Gauker, C. (2011). *Words and images: An essay on the origin of ideas*. Oxford: Oxford University Press.

Gauker, C. (2012). Perception without propositions. In J. Hawthorne & J. Turner (Eds.), *Philosophical perspectives 26: Philosophy of mind* (pp. 19–50). Malden, MA: Wiley-Blackwell.

Gauker, C. (2017). Three kinds of nonconceptual seeing-as. *Review of Philosophy and Psychology, 8,* 763–779.

Glüer, K. (2009). In defence of a doxastic account of experience. *Mind and Language, 24,* 297–327.

Green, E. J. (2019). On the perception of structure. *Noûs, 53,* 564–592.

Gregory, D. (2019). Imagery and possibility. *Noûs*. https://doi.org/10.1111/nous.12275.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology, 160,* 106–154.

Hummel, J. E. (2013). Object recognition. In D. Riesberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 32–46). Oxford: Oxford University Press.

Kind, A. (2016). Imagining under constraints. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 145–159). Oxford: Oxford University Press.

Kind, A. (2018). How imagination gives rise to knowledge. In F. MacPherson & F. Dorsch (Eds.), *Perceptual imagination and perceptual memory* (pp. 227–246). Oxford: Oxford University Press.

Kratzer, A. (1977). What "must" and "can" must and can mean. *Linguistics and Philosophy, 1,* 337–355.

Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2014). A conceptual framework of computations in mid-level vision. *Frontiers in Computational Neuroscience, 8,* 1–19.

Langland-Hassan, P. (2016). On choosing what to imagine. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 61–84). Oxford: Oxford University Press.

Palmer, S. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 259–303). Hillsdale, NJ: Lawrence Erlbaum Associates.

Quilty-Dunn, J. (2019). Perceptual pluralism. *Noûs*. https://doi.org/10.1111/nous.12285.

Siegel, S. (2006). Which properties are represented in perception? In T. Szabó Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford: Oxford University Press.

Singh, M., & Hoffman, D. D. (2001). Part-based representations of visual shape and implications for visual cognition. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Grouping and segmentation in vision* (pp. 401–459). Amsterdam: Elsevier Science.