Why We Love Pictures (for the Wrong Reasons)

A lesson from the picture of a black hole

Lorenzo Sartori

August 26, 2024

Paper accepted to the 2024 edition of the Philosophy of Science Association (New Orleans) and for publication in the related special issue in *Philosophy of science*. This version of the article has been accepted for publication after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections.

Abstract

In this paper, I first show that similarity accounts of scientific pictures fail with more realistic cases of scientific pictures. My primary case study is the picture of a black hole, from which I develop an interpretation-based account of picture representation analogous to how models represent: a picture represents a designated target system iff, once interpreted, it exemplifies properties that are then imputed to the target via a de-idealising function. Then, I show that justification of the inferences from pictures crucially depends on their causal mechanisms of production, in contrast with the standard justificatory strategies we employ for model inferences.

Wordcount: 4386 words (bibliography excluded) and one picture.

Keywords: pictures, scientific representation, justification, black hole

Contents

1 Introduction

Pictures are ubiquitous in science. Astronomers study pictures shot by telescopes and probes to understand how stars form and dissolve, medics use X-ray pictures and MRI scans to detect diseases and provide their diagnoses, and epidemiologists create heatmaps to explain and predict virus spreading patterns. A question then arises about the epistemic function of pictures in science: how do we learn from them about the portions of the world they are meant to represent?

In section 2, I explore the proposal by Meynell (2013) to understand pictures' representation in terms of similarity, and I argue that this is the wrong way to go via a case study, namely, the picture of the black hole M87*. In section 3, I lay down the basic information about how that picture was produced. In section 4, I develop an interpretation-based account of the picture of M87* showing that it represents its target in the sense expressed by the DEKI account of scientific representation (Frigg and Nguyen 2020). Finally, in section 5 I show that while pictures and models are both representations in DEKI's terms, they substantially differ when it comes to the justification of the inferences, we carry out from them about their target systems. Specifically, our inferences from picture like the one of M87* crucially depends on its history of production, in contrast with most examples of scientific models.

2 The Mirage of Similarity

In the literature on depiction, both in aesthetics and the theory of images, there is an important philosophical tradition which focuses on the concept of similarity in order to explain how pictures represent the real world (Wollheim 1987, Hyman 2006, 2012, Peacocke 1987). The application of the similarity view to *scientific* pictures, however, has remained importantly unexplored by the proponents of the similarity view in the context of scientific models (e.g., Giere 2004, 2010 and Weisberg 2013, Ch. 8). An important exception is Meynell (2013), who explicitly wants to clarify the role of similarity in the use of pictures and visual representations in science. Meynell's is for now the best attempt to make sense of the use of pictures in science by appealing to the notion of similarity, so I take her view as a point of reference for my critical analysis of the similarity account in this context.

Her account is built through a combination of, on the one hand, an attack to Perini's (2005) attempt to apply Goodman's (1976) conventionalism to scientific pictures, and, on the other, a constructive proposal inspired by the work of Willats (1997) in psychology. While Meynell accepts that Perini's Goodmanian approach can work well with linguistic or quasi-linguistic visual representations, like schematic diagrams, it remains insufficient for "dense" pictures, like photographs, scans, microscopic and astronomic pictures, and so on. Then, she argues that it is better to understand the use of scientific pictures by employing a similarity view, combined with our knowledge of psychology, theory of perception, and a combination of geometry and optics.

Following Willats, Meynell (2013, 338) argues that the relation between a picture and its target comes in two steps. First, what she calls the pictures primitives (lines, points, coloured areas) are associated with scene primitives, namely, the most elementary units of shape information in the scene – which can be 3D (lumps, sticks and slabs), 2D (surfaces), 1D (edges) or 0D (corners). Second, these scene elements are then related to the target system in the real world.

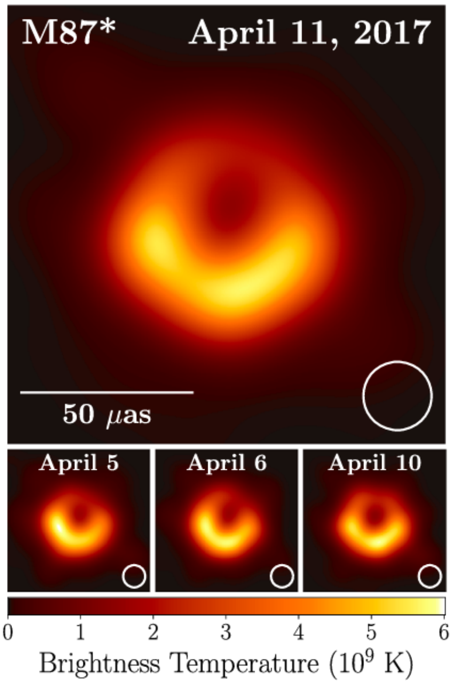

Figure 1: Picture of the black hole M87*. In EHTC (2019a, p. 5).

Meynell then characterises both the relation between picture primitives and scene primitives, and the relation between scene primitives and the target. The aim is to show that both steps are based on (objective) similarity and how our perception works, in contrast with the position defended by Perini and Goodman, who focus on interpretation and convention. The first step, from picture primitives to scene primitives, is expressed in terms of geometrical projection: the picture primitives are just the result of geometrically projecting the elements of the scene according to a specific system of projection (say, perspectival, orthogonal, or oblique). The second step, which connects the scene with the actual target system, will be instead mediated by visual or perceptual similarities: the scene represents the target insofar as when we look at the scene, we see something similar to what we would see if we were looking at the target system itself.

Let us assume that Meynell's account succeeds in explaining how photographs, realistic paintings, and simple geometrical figures represent. I want to offer a new argument against the similarity view, in addition to more general considerations that have been already covered in the literature. The point I want to make here is that even if the similarity view were successful for more mundane cases, the account is inadequate when applied to more complex examples of visual representations used in science.

For my argument, I will mostly focus on the recent picture of the M87* black hole at the centre of the Messier galaxy produced between 2017 and 2018 in the context of the project Event Horizon Telescope (figure 1). Let us consider this picture for a moment.

First, according to Meynell, we should be able to see the scene represented by the picture by geometrically projecting the picture primitives (in this case, coloured regions) into the scene basic elements. However, even if we know what geometrical projection is in place (in this case, a form of perspectival projection), it is unclear what the scene would be. This is because geometrical projections are neutral with respect to what is projected. Second, if we take that the similarity between the scene and the target is a matter of visual appearance, we end up inferring that the black hole is reddish-yellow doughnut in the middle of nothing, which is strikingly incorrect.

What we need, instead, is a function that associates colours with what these colours are in fact meant to represents: levels of radiation intensity of an electromagnetic field. In other words, we need an interpretation of the picture that translates visual properties into another set of non-visual properties. So, on closer inspection, similarity is not really the essential concept on which to base the epistemic use of pictures as representation, the lion share being played by such an interpretation. Once the translation of colours in levels of radiation intensity is in place, there is no meaningful sense in which the picture and the target are similar, particularly if similarity is understood in visual, perceptual terms.

This point can be generalised by looking at other examples of scientific pictures, drawing examples from Elkins (1999) and Tufte (1997).

3 One step back: imaging a black hole

There is already an emerging literature on the epistemology of the picture of M87* and in general of black holes (Skulberg 2021; Muhr 2023; Doboszewski and Lehmkuhl 2023; Doboszewski and Elder 2024). Curiously enough, however, philosophers of science have not studied this picture *as a picture*. That is, they have not focused on the features of this picture as a representation, namely as an object allowing surrogative reasoning about its target system. I want to suggest that, while the epistemological analyses conducted so far are crucial, they remain incomplete. Indeed, a study of the picture as evidence presupposes an analysis of how the picture of a black hole is supposed to be "read" as a representation. In this sense, my analysis will also be a useful contribution to the general epistemological enquiries about black hole pictures.

In order to make such an analysis of the picture of M87* as a representation, one needs to first provide some details on black holes and how we image them.

Black holes are astronomic bodies so dense that they gravitationally trap anything comes close enough to them, light included. Therefore, they are in principle invisible. However, given their exceptional gravitational pull, they also produce a vortex of matter, mostly ionised gases, that orbits around them –which is called *accretion disk*. Our observations are meant to study the shape of this disk in order to understand basic features of the back hole and to test the predictions of general theory of relativity.

The fundamental theory to measure astronomic phenomena is interferometry. An interferometer gives a measure of the intensity of radiation of an electromagnetic source by decomposing the original light in two beams and then calculating the phase differences between them. In this way, we can measure the relevant distribution of the radiation from a radiating source – in principle, even from an astronomic one like the accretion disk of the black hole M87*. These measurements are crucial, because the distribution of radiation can give us insight on the dimensions and shapes of what lies within the event horizon, which is usually called the *shadow* of a black hole.

For four days in April 2017, seven telescopes in different locations on the globe were pointed towards the centre of the Messier 87 galaxy and measured the radio signals coming from that region of spacetime. The idea was to synchronise all the telescopes so that they could be used as one single telescope. The resulting "lens" of this composite telescope, even though fragmented, had the width of the entire planet Earth. One and half petabytes of interferometric data were collected for each night of observation, that is, the greatest amount of data in the history of science for single experimental measurement. These data were then fed to a supercomputer which integrated the data of each single telescope. The data was further calibrated (the Event Horizon Telescope Collaboration (EHTC) *et al.* 2019b).

Then, four teams of researchers were created to independently produce a visual output from the data. The four teams worked autonomously, and they were not allowed to talk with each other. Two teams ended up using the so-called CLEAN algorithm, while the other two teams used two different versions of the so-called Regularized Maximum Likelihood (RML) family of algorithms: the algorithm SMILI and the algorithm EHT-imaging, the latter created specifically in the context of the Event Horizon Telescope measurement (EHTC *et al.* 2019c).

Then, the four teams compared their output images, which all exhibited two important structural features: a ring shape with more intense brightness in the south region of the black hole, and the diameter of the ring estimated around 40μas (*ibid.*, 9). Four images were then produced from each algorithm pipeline, one for each night of observation. As a further step in making these pictures more reliable, all these four images were further blurred to obtain a "common, conservative resolution" of each of them (*ibid.*, 20). Finally, to further emphasise the common features of the images produced by the three different pipelines, the scientists produced an average picture for each of the four days (*ibid.*, 21).[1]

The image that we eventually obtain from this procedure is, basically, a heatmap. A heatmap is a visual representation of data where values are represented by colours, and the spatial coordinates on the map are to be translated in other properties of the represented phenomenon. The easiest case is when the spatial properties on the map are translated in spatial properties of the target system via a geometrical projected. However, heatmaps can use colours to represent any sort of property or quantity. What we need is a way to systematically interpret the visual properties of the heatmap into the relevant properties we are actually measuring in the target system. In section 4, I draw an account of how this works with the picture of the black hole M87*.

4 An interpretation-based account of scientific pictures

The picture of M87*, I argue, interpreted as such, refers to M87*. Following Goodman (1976), I take this referential relation from a symbol to an object to be denotation, namely, the referential relation between a name and its bearer. In order to denote M87*, though, we need to look at the picture as not, say, mere pixels coloured in a specific way: we need to interpret the material instantiation, or *carrier*, of the picture, as a radiation-heatmap of the electromagnetic field surrounding M87*. Interpretation here can be understood as a function I that maps colours of the picture into levels of radiation intensity of an electromagnetic field, where black is interpreted as lower levels and yellow as higher. So, we have three elements: the carrier, the heatmap, and the target. The picture, once interpreted, represents M87* as a radiation-heatmap – or, alternatively, it is radiation-heatmap-representation of M87*.

The heatmap resulting from our interpretation of the carrier is not supposed to be read as a truthful description of the target, as it may involve idealisations and distortions

[1] 1 This is of course a very simple reconstruction. The reader can find all the details in the six articles published by the EHTC team reported in the bibliography.

with respect to the target system. Again, I use Goodman's (1976) terminology and I call the radiation-heatmap a Z-representation. Distinguishing these three elements (carrier, Z-representation, and target system) is crucial to appreciate the type of reasoning involved when we make inferences about the actual black hole on the basis of the picture. This reasoning is surrogative in nature (Swoyer 1991): reading the picture as a representation requires us to study a system in order to make inferences about another system.

It is nevertheless useful to study the Z-representation in question because it highlights certain properties of the target that we would have not been able to detect if we had just looked at the raw interferometric data collected by our telescope. Using the terminology of Goodman's (1976) and Elgin (1983, 1996) we can call this function *exemplification*. Technically, an object exemplifies a property A if it instantiates A and refers to A by instantiating it. A typical example is a swatch in a tailor shop. The swatch possesses many properties (say, being rectangular, being produced in Thailand, weighing 1.5 grams...), but, in the context of the shop where people decide how their clothes should be, the swatch refers only to a certain set of properties: colour, texture, material, and so on. By referring to them, it makes them salient: these properties become, or should become, more epistemically accessible for a user or observer.

The same happens with the radiation-heatmap: by abstracting away informational noise and resulting by an interpolation of the original sparse data, this picture summarises and highlights certain prominent features of M87*, particularly, the remarkable shadow of the black hole, its dimensions, and an asymmetry in radiation intensity between the Southern and Northern areas of the accretion disk.

The point of all this is of course to eventually impute some of these exemplified properties to the target system, the actual black hole M87*. When I talk about imputation, I simply mean property attribution, with no assumptions about whether this attribution is correct or not.

Sometimes, the properties exemplified by a representation can be imputed to its target system unchanged. Sometimes, however, this imputation will require some form of de-idealisation, a further interpretive activity. To distinguish this step from the interpretation function I, let us call key the function that translates idealised properties of the representation into non-idealised properties imputed to the target.

In the case of the picture of M87*, there are many of these keys at work. One is simply a scale factor, multiplying the dimensions of the object in the picture into the actual dimensions that we expect M87* has, based on the picture together with our knowledge of the distance between us and the centre of the Messier galaxy. Another important key is a geometrical projection that translates the two-dimensional spatial properties of the heatmap into a set of three-dimensional ones.[2] A final interesting type of key seems in place in the process of blurring the final visual outputs of the algorithms in order to decrease precision but increase reliability. Here, a key should be used to de-blur the picture if our aim is to attribute more fine-grained, precise properties to the actual black hole.

From this reconstruction, one can see that no appeal to similarity, perception or psychology has been made: given the presence of the I and the key, neither the picture as a carrier nor the picture as heatmap need to be similar to the target to represent it. At the same time,

[2] This geometrical translation will have to account for the complex geometry of the spatiotemporal region under investigation, as distances between areas in the pictures are distorted with respect to actual distances between the corresponding regions of the black hole.

the account still retains Meynell's basic intuition about the role of geometrical projections in visual representation. This role, though, is inserted in a more general framework entirely based on interpretation (in the various fashions of the I-function, denotation, the selective process of exemplification, and the key). Interpretation here is arbitrary but not random: while heavily theory-laden, our interpretation of the picture strictly depends on the way in which the project was in fact produced.

Once the picture is interpreted as a heatmap, four basic elements have been highlighted in my analysis: denotation, exemplification, keying-up, and imputation. These are the basic ingredients of the so-called DEKI account of scientific representation (Frigg and Nguyen 2020), that takes its name exactly from those ingredients. More formally, according to DEKI, a model system M (that is, a carrier C endowed with an interpretation I) is an epistemic representation of a designated target system T iff four conditions apply:

(i) M *denotes* T,

(ii) M *exemplifies* properties P_1..., P_n,

(iii) P_1..., P_n are associated with a second set of properties Q_1..., Q_n via a *key*,

(iv) Q_1..., Q_n are *imputed* to T.[3]

I can then conclude that, at this level of analysis, a mechanically produced picture like the picture of M87* function as an epistemic representation in the same way as scientific models represent their target systems.

5 From semantics to epistemology: measurement vs. model

So far, I have focused on how to interpret the picture as a representation, and I have characterised this representational function in the terms of the DEKI account. However, the account has two important shortcomings. First, it is skeletal by design: it needs to be completed with the specifics of each case study. The analysis of the black hole picture just offered provides the relevant details on how to apply DEKI to this specific case study. Second, the account remains silent on the justification of our inferences from the picture to the target system. The account highlights the elements constituting our reasoning process that from the representation allows to make hypotheses about the target system, but it does not say how to assess the reliability of these inferences. This is because the account correctly acknowledges that the justificatory roots of our inferences lie outside the single representation system. In this respect, pictures are exactly like models, because even for inferences drawn from a model about a target system, the only way to justify our inferences is to support them with reasons extrinsic to the single model system (theories, observations and other models).

The similarity between pictures and models, though, ends here. For, I want to suggest, in the case of pictures like the picture of M87* the root of the justification of both our interpretation of the picture and of our inferences about its target is the same, namely, the

[3] The DEKI account is very complex and an exhaustive analysis of it lies outside the scope of this paper. Interested readers can find all the details in Frigg and Nguyen (2020, 159-214).

causal history of production connecting a given picture to its designated target system. Let us assume that, like in this case, it is impossible to observe the target system directly. If one wants to convince me that the picture is giving me (approximately) correct results about the actual black hole, they will have to explain to me how the picture was produced, and how the visual output we are looking at is causally dependent on the target system.

For example, Doboszewski and Elder (2024) analyse the picture in terms of robustness analysis, by showing that the multiple algorithms employed for the imaging converged on similar results even if taking different procedures and assumptions, and they also exhibited reasonable sensitivity to data – namely, if the data had been different, the resulting visual output would have been different in a consistent, systematic manner. The robustness of the algorithms was assessed as follows. The three algorithms were tested against synthetic images, showing different geometrical shapes, which the algorithms had to reconstruct as with the picture of the black hole. This was done by surveying a broad range of combinations of parameter values. By this parameter survey on synthetic images, the researchers obtained two results. First, they identified the fiducial parameters, that is, those parameters that allowed a more faithful reconstruction of the original image. Second, they proved some robustness of the algorithms by showing that they were sensitive to the input image: the outputs were really different for each synthetic image and the black hole picture, showing that there was a relatively strong counterfactual dependence of the visual output on the original source.

This was necessary to secure a reliable counterfactual relation between the data and the visual output obtained by applying the algorithm, and consequently, the accuracy of the latter with respect to the former.

The causal relation that I suggest lies at the root of the inferential stability[4] from pictures to targets more generally, and thus the reliability of the former as successful representations of the latter. The more numerous and complex are the steps in the causal chain of producing the picture from the target, the more difficult it will be to justify the inferences we draw from the former to the latter. Here, I am not arguing that the picture of M87* is, in fact, epistemically reliable. My point is just that if one wants to assess such reliability, one has to look at its production, and how this relates with the interpretation of the picture and the de-idealising keys employed.

Again, it is important to notice that the notions of reliability, accuracy, and success that I am employing here do well without any appeal to similarity between the picture and the target system: what counts is the counterfactual stability between representation and target, which in turns depends on the causal mechanisms connecting them.

Nothing of the sort of what I have said about justification in the case of the picture of M87* applies to other forms of representations, like models. A model system is usually constituted by a set of assumptions on that system (an abstract object or a material one), often in interaction with each other. Let us take the simple case of an assumption that is expressed by a certain functional relation between two quantities. There are many ways in which we can justify this assumption. It may directly derive from more general theory in the relevant discipline. Or our assumption may be a simplification of a more general functional relation that however is intractable in its current form (e.g., an equation with no analytic solutions). Here, the justification follows from our reason to hold the original formula, plus some further reason to consider the simplification acceptable. Alternatively, our assumption

[4] Cf. Roskies (2008) for similar considerations applied to MRI scans.

may boil down to a hypothesis abstracted away from data. For example, it may be the result of an abductive inference on the basis of current observations. Here, some further justificatory analysis is required for the inference to the best explanation.

In all these cases, the justification of the assumption will be more or less provided on the basis of previously acquired knowledge. However, the assumption could also be something completely new, detached from theory and experiments. The justification of that assumption will then solely depend on the success of the model as a whole. Success can take many forms: empirical adequacy, unification, explanation by providing an underlying mechanism. The more the model proves itself successful, the more we can justify its further application as an epistemic surrogate system. However, as it should be evident, there is no reference to causal relations between the target and the model in our justification of the inferences from the latter to the former.

Nevertheless, it is important to remind the reader that the characterisation of this causal relation is still based on theoretical assumptions and previously acquired empirical knowledge. I do not want thus to undermine the theory-ladenness of our interpretations of images: whether an image is causally linked to its target and how accurately so can, and is, a matter of dispute, even among experts. So, I do not want my focus on causation here to foster the suspicion that I am considering pictures somewhat more "objective" representations than models or other types of representations. For even the assessment of the hypotheses about the causal relations in play in our production of images will strictly depend on the theoretical framework we are assuming in the first place.

6 Conclusion

In this paper, I argued that similarity accounts of pictures fail with more complex scientific cases like the picture of the black hole M87*. I then offered an interpretation-based account of this picture, where the interpretation of the picture was properly anchored to the way in which the picture was produced. I show that, *qua* representation, this picture does not differ from other types of representations like scientific models. The difference, instead, concerns the justification of the inferences we draw about the designated target. In the case of pictures like M87*, the inferences can be justified only by an appeal to the causal process of production of the picture, while this is not a common justificatory strategy in the case of scientific models.

References

Doboszewski, J., & Elder, J. (2024). Robustness and the event horizon telescope: The case of the first image of M87. *arXiv preprint arXiv:2401.16323.*

Doboszewski, J., & Lehmkuhl, D. (2023). On the epistemology of observational black hole astrophysics. In N. M. e. a. Boyd (Ed.), *Philosophy of astrophysics: Stars, simulations, and the struggle to determine what is out there* (pp. 225–247). Springer.

Elgin, C. Z. (1983). *With reference to reference.* Hackett Publishing.

Elgin, C. Z. (1996). *Considered judgement.* Princeton University Press.

Elkins, J. (1999). *The domain of images.* Cornell University Press.

Event Horizon Telescope Collaboration, *et al.* (2019a). First M87 event horizon telescope results. I. The shadow of the supermassive black hole. *The Astrophysical Journal Letters*, *875*(1), L1.

Event Horizon Telescope Collaboration, *et al.* (2019b). First M87 event horizon telescope results. III. Data processing and calibration. *The Astrophysical Journal Letters*, *875*(1), L3.

Event Horizon Telescope Collaboration, *et al.* (2019c). First M87 event horizon telescope results. IV. Imaging the central supermassive black hole. *The Astrophysical Journal Letters*, *875*(1), L4.

Frigg, R., & Nguyen, J. (2020). *Modelling nature: An opinionated introduction to scientific representation*. Springer.

Giere, R. N. (2004). How models are used to represent reality. *Philosophy of Science*, *71*(5), 742–752.

Giere, R. N. (2010). An agent-based conception of models and scientific representation. *Synthese*, *172*(269), 269–281.

Goodman, N. (1976). *Languages of art*. Hackett.

Hyman, J. (2006). *The objective eye: Color, form, and reality in the theory of art*. University of Chicago Press.

Hyman, J. (2012). Depiction. *Royal Institute of Philosophy Supplement*, *71*, 129–150.

Meynell, L. (2013). Parsing pictures: On analyzing the content of images in science. *The Knowledge Engineering Review*, *28*(3), 327–345.

Muhr, P. (2023). The "cartographic impulse" and its epistemic gains in the process of iteratively mapping M87's black hole. *Media & Environment*, *5*(1), 341–357.

Peacocke, C. (1987). Depiction. *The Philosophical Review*, *96*(3), 383–410.

Perini, L. (2005). The truth in pictures. *Philosophy of Science*, *72*(1), 262–285.

Roskies, A. L. (2008). Neuroimaging and inferential distance. *Neuroethics*, *1*, 19–33.

Skulberg, E. (2021). *The event horizon as a vanishing point: A history of the first image of a black hole shadow from observation* (Ph.D. Dissertation). University of Cambridge.

Swoyer, C. (1991). Structural representation and surrogative reasoning. *Synthese*, *87*, 449–508.

Tufte, E. R. (1997). *Visual and statistical thinking: Displays of evidence for decision making.* CT: Graphic Press.

Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world.* Oxford University Press.

Willats, J. (1997). *Art and representation: New principles in the analysis of pictures.* Princeton University Press.

Wollheim, R. (1987). *Painting as an art.* Thames; Hudson.