# Generative AI and photographic transparency

P.D. Magnus*

November 2, 2023

**Abstract**

There is a history of thinking that photographs provide a special kind of access to the objects depicted in them, beyond the access that would be provided by a painting or drawing. What is included in the photograph does not depend on the photographer's beliefs about what is in front of the camera. This feature leads Kendall Walton to argue that photographs literally allow us to see the objects which appear in them. Current generative algorithms produce images in response to users' text prompts. Depending on the parameters, the output can resemble specific people or things which are named in the prompt. This resemblance does not depend on the user's beliefs, so generated images are in this sense like photographs. Given this parallel, how should we think about AI-generated images?

## 1 Photographic transparency

In "The Will to Believe" (1896) William James reflects on the fact that belief is largely involuntary. He asks: "Can we, by just willing it, believe that Abraham Lincoln's existence is a myth, and that the portraits of him in *McClure's* Magazine are all of some one else?" [7]. When James was writing, Ida Tarbell's biography of Lincoln had recently been published in *McClure's*. One scoop that Tarbell uncovered was a previously unpublished daguerreotype of young Lincoln. The photographs were a key sales point for the magazine, which also included familiar images of older Lincoln [5]. For example, see figure 1. This is why Lincoln was a salient example for James, more so than say George Washington. Although there were paintings and sculptures of Washington, there were photographs of Lincoln.

A painting or sculpture can only include the features of the subject to which the artist was sensitive, but a photograph is mechanical. As Oliver Wendell Holmes had written several years earlier, "In a picture you can find nothing which the artist has not seen before you; but in a perfect photograph there will

*This is the author's final draft of a paper forthcoming in *AI&Society*. The author can be reached via fecundity.com. Thanks to Evan Malone, Dan DiTursi, Jason D'Cruz, and Ron McClamrock for helpful discussion of these issues.

be as many beauties lurking, unobserved, as there are flowers that blush unseen in forests and meadows" [6, p. 744]. A photograph does not rely upon and is not sensitive to the beliefs and attitudes of the photographer, at least not in the way that a painting or drawing is sensitive to the beliefs and attitudes of the artist. It is easier to imagine that the subject of a painted portrait is a myth than the subject of a photographic portrait. This is not because photographs cannot be faked, because of course they can. Rather, there is a sense that photographs present things more directly than paintings and other illustrations. For images which we know are fantastical, we nevertheless use the adjective "photorealistic" to mean how a thing actually would look— rather than how it would typically be drawn or painted.

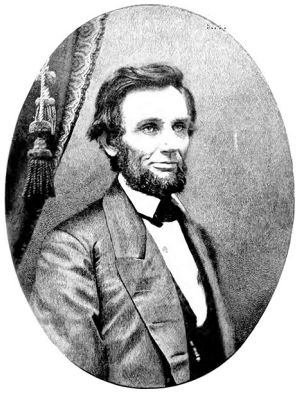

**Figure 1:** An 1861 photograph of Abraham Lincoln, reproduced in Ida Tarbell's *The Life of Abraham Lincoln* [13, p. 101]. This digital image is extracted from a PDF file of a digital scan of the book.

This point about photographs is taken up by Kendall Walton, who writes that "photographs of Abraham Lincoln, for instance, are in some fundamental manner more realistic than painted portraits of him" [14, p. 251]. What matters, he argues, is the way in which Lincoln caused the image in the photograph. Photography is like seeing, in that it is not mediated by an artist's beliefs. He sums the point up this way: "part of what it is to see something is to have visual experiences which are caused by it in a purely mechanical manner. Objects cause their photographs and the visual experiences of viewers mechanically; so we see the objects through the photographs. By contrast, objects cause paintings not mechanically but in a more 'human' way, a way involving the artist; so we don't see through paintings" [14, p. 261]. This conclusion is the claim that photographs are *transparent*, meaning that seeing the photograph of Lincoln is an instance of actually seeing Lincoln.

Of course there is a gap in this inference. Walton only claims to show that being mechanical in this way is *part* of what it is to see something; that is, it is a necessary condition for seeing. There might be other necessary conditions for seeing which photography lacks, conditions that would block transparency. Indeed, critics of Walton have pointed out differences between, for example, my seeing a coffee cup (on the table in front of me) and my seeing Abraham Lincoln (in the photograph). Gregory Currie [4, ch. 2] and Noël Carroll [1] note that when one sees something in the ordinary way, one can situate it in one's

egocentric space. When I see the coffee cup, I see it as being in front of me. When I see the photograph of Abraham Lincoln, I see the photograph as being on the screen in front of me but cannot situate Lincoln or the scene he occupies in relation to my surroundings. This argument has been refined by Jonathan Cohen and Aaron Meskin [3, 9] and by Bence Nanay [11].

The critics are right, of course, that there are some differences between my awareness of the coffee cup (which we might call *direct seeing*) and my awareness of Abraham Lincoln (which we might call *seeing in a photograph*). I am unsure whether the unqualified word "seeing" describes the latter, but that is primarily a question about the word. Even if the critics are right that I do not *see* Lincoln in an unqualified sense, Walton is still right that there are similarities between the two cases which distinguish them from my awareness of objects depicted in a painting. Part of what direct seeing and seeing in a photograph have in common is the feature highlighted by Walton (and before him Holmes) that neither is mediated by an artist's beliefs about the object. With respect to the divide between direct seeing and seeing in a painting, seeing in a photograph is on the side with direct seeing. A photograph, as Walton says, is mechanical.[1]

It is worth noting two extensions of this point before moving on.

First: Writing in the 1980s, Walton did not anticipate digital photography. Yet, now that we have smartphones in our pockets, most photographs we encounter are digital.[2] Regardless, his argument for transparency naturally extends to digital photographs. Vivian Mizrahi [10] makes the extension explicit. In both cases, she writes, "the chromatic discontinuities on a photograph are the end result of a complex causal process that relies on the fact that the multitude of points— or pixels— constituting the surface of the photograph are individually correlated to properties of the light rays selected by the lens" [10, p. 59].[3]

Second: The relation of mechanical presentation is transitive. The picture of young Lincoln in *McClure's* was not a mere photograph, but instead "a halftone reproduction of a photograph of a daguerreotype" [5, p. 38]. If one can see Lincoln in the original daguerreotype, though, one can similarly see him in the photograph of the daguerreotype, the halftone reproduction of the photograph, and a digital scan of the reproduction.

[1] In later work, Walton suggests that the word *mechanical* in this context is "far from clear" [15]. He writes instead about the *photographic relation* and expresses his view in this way: "[T]he feature of photographs that I took to be crucial to their transparency is key to understanding the photographic relation: Photographs depend counterfactually on the objects photographed, even when the beliefs and other mental states of the photographer (and anyone else) are held fixed" [15]. When I use the word *mechanical*, I use it to mean this crucial feature.

[2] Christy Mag Uidhir [8, fn. 7, p. 40] argues that film photography should be understood under the genus of printmaking and therefore that so-called digital photographs are not really photographs at all. Even so, there is a larger category of images that includes old-school photographs and digital photographs. My use of the word "photograph" in this paper refers to that larger category.

[3] In later work, Walton is explicit that his account should also apply to pictures taken with a digital camera [15].

## 2 Generative AI

What should we say about other digital images?

A digital drawing or illustration, entered with a mouse or stylus, is clearly on the side with paintings. Features of the objects depicted are mediated by the operator's beliefs.

Some digital photographs are filtered or supplemented with extra graphics, selected in real time by the person using the camera. Insofar as their selection of extras is guided by their beliefs about what it is that they are photographing, the resulting images might have at least some of the human quality of illustrations.

There are now, however, a great many digital images that are generated without that kind of mediation. Algorithms, given a text prompt and some other parameters, can generate a digital image or collection of images. Current examples of such systems include Dall-E, MidJourney, and Stable Diffusion. Setting aside generative algorithms which produce other kinds of outputs, like large language models that produce text, I will refer to the image generation algorithms simply as *generative AI*.

One might worry that generative AI threatens the connection between the image and its object. Consider figure 2, the result of a prompt for *Abraham Lincoln in a funny hat*. The hat is rather disappointing, looking like nothing more than a pom pom or glow on the back of his head. Even so, there is no historical evidence that suggests Lincoln ever wore such a pom pom.

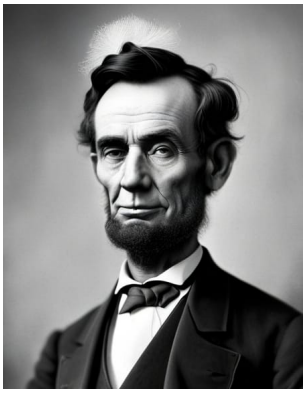

**Figure 2:** An image generated from the prompt, "Abraham Lincoln in a funny hat; portrait; historically accurate" using a version of the Stable Diffusion algorithm at nightcafe.studio

We might take this to be a benign example of what worries Regini Rini in her discussion of deepfakes [12]. Her primary focus is audio or video recordings that seem to show a well-known person doing something they never did, but she begins with photographs that are manipulated or falsified. Drawing on work by Dan Cavedon-Taylor [2], she holds that photographs have two important features.

First, "when we see a photograph that depicts $x$ as $F$, say, our default doxastic response is to believe that $x$ is $F$ — and to only withhold assent if we possess reasons against thinking the photograph creditworthy" [2, p. 294][12,

p. 9]. That is, when we see features in a photograph, our default rule of evidence is to believe them.

Second, seeing a photograph allows us to form perceptual knowledge, while seeing a handmade picture (like a painting) only allows for testimonial knowledge. The photograph can produce new knowledge (letting the viewer learn something the photographer did not know), whereas the painting can only transmit knowledge (letting the viewer learn something that the painter knew well enough to paint).

For Cavedon-Taylor and Rini, these two points are related— "perceptual evidence... enjoys a stronger presumptive authority than testimonial evidence" [12, p. 10]. For a painting, the viewer must rely "upon the painter's skill and honesty" [12, p. 9]. They allow, of course, that photographs can be edited or fabricated. When that concern is salient, we should not believe as a default whatever we see in photographs. If any would-be photograph might just as well be a forgery or the product of a generative AI, Rini suggests, then gaining knowledge from a photograph also depends on the honesty of the person presenting the image. Rini writes, "Perhaps the evidential status of still photos has *already* shifted from vehicles of perceptual knowledge to mere testimonial knowledge" [12, p. 13].

This conflates the two points. Although related, they are not equivalent.

The first feature does not entail the second: The policy of taking photographic evidence that $x$ is $F$ as presumptive reason for thinking that $x$ is $F$ is sensible if photographs are typically veridical. And the same condition might motivate a similar policy for certain drawings. If official blueprints are almost always accurate, then it would be a good policy to believe that $x$ is $F$ on the basis of an official blueprint that depicts $x$ as $F$. Yet the blueprint is still testimonial (in Cavedon-Taylor's sense). The depiction is mediated by the attitudes and beliefs of the drafter. As Walton would put it, the objects are related to the blueprint in a human way rather than in a mechanical way.

The lack of the first feature does not entail the lack of the second, either: There are assumptions we need to make in order to trust a photograph— as Walton writes, we must assume "that no monkey business was involved in the processing" [14, p. 263]. Moreover, we need to believe that the image we are looking at is a photograph, rather than a photorealistic painting. These assumptions might be unjustified. If someone presents me with an image which they claim to be an unedited photograph, then I may have to trust them. Importantly, this does not transform the photograph itself into testimonial evidence; consider two reasons: First, rather than merely trusting the person who presents the would-be photograph, I might investigate the chain of evidence for myself. I might uncover perceptual and inferential reasons to think that the image is a photograph and that nobody had tampered with it. Second, if the image is in fact a photograph, then it has precisely the mechanical quality that Walton highlights. If I see in the photograph that $x$ is $F$, my belief that $x$ is $F$ may depend on testimony that *this is a photograph*— but there is no need for testimony about $F$ and whether it is $x$.

My interest here is in the mechanical nature of photographs and in the

perceptual nature of our encounters with them. That is independent of whether it is wise to presume that would-be photographs provide knowledge, so I set that epistemic concern aside.[4]

## 3 How transparent are generated images?

Imagine I were to take a photograph of a funny hat, cut it out, and paste it onto a photograph of Lincoln. Someone who looked at the composite might end up believing that Abraham Lincoln wore a funny hat, but such a belief would not just result from seeing a photograph of Lincoln. Rather, it would result from seeing two photographs artfully juxtaposed.[5] Importantly, the Lincoln bit of the composite and the hat bit of the composite would each, separately, have the mechanical features that make seeing in a photograph more like perception than like testimony.

So return now to figure 2, the AI-generated image which seems to show Abraham Lincoln with a pom pom on his head. Although the image is not a photograph, it does correctly represent the course-grained appearance of Abraham Lincoln. Moreover, the features of Lincoln which appear in the image do not depend on my beliefs and attitudes. I entered a prompt, and the algorithm did the rest. So the image is mechanical.

This is not to say that figure 2 is a photograph. Rather: It has the mechanical nature which photographs have and which paintings lack. This feature is plausibly a necessary feature of perception. As Walton's critics are quick to point out, it is not a sufficient condition. Even so, this result puts images like figure 2 on the side with perception and photography opposite to painting and drawing. For all that has been said so far, seeing Lincoln in a generated image (like figure 2) is the same sort of thing as seeing Lincoln in a photograph (like figure 1).

I do not think that this result is correct, but it is harder to rebut than you might expect. I will consider several possible replies, only the last of which is decisively successful.

First, one might argue that figure 2 is a mere picture of Lincoln— that is, it depicts Lincoln in only the way that (for example) the stamped face of a penny does. I do not have a general account of *depiction*, so I have nothing deep to say on this point. Regardless, being a depiction does not preclude also having the mechanical quality that typifies photographs. Many photographs also depict

[4] My engagement here with Rini is just to resist her conflation of the two issues. Her primary focus is on the erosion of epistemic authority; that is, what happens when it becomes unwise to presume that recordings provide knowledge. Addressing that is beyond the scope of this paper.

[5] Walton would say that the viewer does not see Lincoln wearing a funny hat but *does* see both Lincoln and a funny hat. See his parallel discussion of photographic composites made in a darkroom from multiple negatives [14, p. 268–9]. Cavedon-Taylor would say that the resulting composite is not "strictly speaking, a photograph" but instead a "photomontage" [2, p. 288].

things.[6] It is just the question of the mechanical quality that interests me here— a quality that figure 1 and figure 2 share— so let's set aside the question of depiction.

Second, one might focus on how the algorithm is used. A typical way to use generative AI is iterative. The user starts with a prompt and generates several output images. They then select one of those to use as a new input to the algorithm, possibly revising the prompt as well. This leads to further outputs, which can either be accepted as good enough or used as the basis for further runs. The user decides how to proceed at each step, so the final image ultimately relies on their beliefs and attitudes. Although it is not exactly like a painting, the product of such an iterative process is not mechanical in the way a single photograph is. Regardless, that is not how I generated figure 2. Instead, figure 2 is the first result generated by the algorithm given my initial prompt.

So one might instead try to exploit the fact that I did not do much to generate figure 2. Rather, most of the work was done by the generative AI. If we consider the programmers and engineers who made the AI, their beliefs and attitudes influenced the resulting image. However, this is only true for their beliefs and attitudes about algorithms, data sets, computing resources, and what not. Their beliefs and attitudes about Lincoln played no more of a role than mine did.

Third, one might try to deny that the image in figure 2 counterfactually depends on Lincoln's visage. This strikes me as implausible. Admittedly, the causal history of the image is rather more complicated than the causal history of a photograph, and there is no way to draw a one-to-one map between pixels in the image and points on Lincoln's face at a particular time. Nevertheless, if Lincoln had looked different, then he would probably appear different in the image. If Lincoln had lost his nose as a young boy, for example, then the nose in the image would look different. This is because the algorithm's training set includes photographs of the historical Lincoln, and those photographs mechanically document how his nose looked. This just underscores the parallel between the algorithm and photography: Just as a photograph of a photograph has the mechanical quality that Walton highlights, the image generated by the algorithm trained on photographs has that mechanical quality.[7]

Fourth, note that the training set for the algorithm was not just an undifferentiated collection of images. Rather, each of the images was tagged with words or phrases. The algorithm came to associate the name "Abraham Lincoln" with something like the visage presented in figure 2. The facts that the data set had to be tagged and that the image is generated in response to a text prompt provide the clue to a crucial difference between generative AI and photography.

[6] See Walton [15] on the distinction between the depictive relation and the photographic relation.

[7] One might press the objection by noting that the training set for the algorithm probably includes paintings and illustrations of Lincoln as well as photographs. Those contribute in some small way to the output and thus introduce a small human element in figure 2, making it less than fully mechanical. Regardless, we can imagine— one could even construct— a generative AI trained only on photographs. The conclusion that the looking at the output of such a generative AI would be like seeing in a photograph remains odd.

Interestingly, it is one that Walton himself (almost) anticipates.

Recall that Walton thinks that the mechanical nature of photography is necessary for it to count as perception. This is the opening that allows his critics to suggest other necessary conditions, ones which photography does not meet. Yet Walton himself proposes a further necessary condition for seeing, one which has gone mostly undiscussed in the literature.

Walton imagines a machine which detects incoming light and prints out a text description of the scene. He thinks it is obvious that reading the output of such a device would not count as seeing the scene. However, the text descriptions would not be mediated by someone else's beliefs (the way a painting is) so there must be some further feature required for transparency— a feature that photographs have but which the machine-generated text descriptions would lack. This feature (he suggests) is that the kinds of mistakes we are inclined to make with photographs are like those that we make with ordinary direct seeing, but the kind of mistakes we make with text descriptions are entirely different. He writes: "A *house* is easily confused with a *horse* or a *hearse*, when our information comes from a verbal description, as is a *cat* with a *cot*... and so on. When we confront things directly or via pictures, houses are more apt to be confused with barns or woodsheds, cats with puppies, and so forth" [14, p. 270].

The machine that Walton imagined now exists, but it is roughly the reverse of a generative AI. Instead of taking an image and returning a text description, a generative AI takes text and generates an image. Regardless, the fact that it has text at one end makes it prone to the sorts of errors that Walton highlights. It trips up on homophones— for example, depicting fruit when prompted for an orange color scheme. And things go wrong when the prompt is not spelled correctly.

As evidence, consider figure 3. The prompt includes the deliberately obscure homophone "Ehbruham Lynkon", and the output is not particularly like Lincoln at all. (The funny hat is rather more conspicuous than in figure 2, however.)

**Figure 3:** An image generated from the prompt, "Ehbruham Lynkon in a funny hat; portrait; historically accurate" using a version of the Stable Diffusion algorithm at nightcafe.studio

Whether a photograph (like figure 1) or an AI-generated image (like figure 2) can provide knowledge about Lincoln depends on the causal chain of evidence, the reliability of the counterfactual dependence of features in the image on facts

about Lincoln. Walton's further criterion draws our attention to a different range of counterfactuals: We ask not how the image would have been different if Lincoln had been different, but instead what kind of confusions were likely to have resulted if things had gone wrong. A photograph would give rise to the kinds of confusions that direct seeing would, mistaking something for another visibly similar thing. A generated image allows those kinds of errors but also allows mistakes which turn on verbal or textual similarity.

# 4 Conclusion

Walton's first criterion, mechanical presentation, groups seeing in a photograph together with direct seeing. It distinguishes those from seeing in a painting or drawing. As I have argued here, this criterion groups AI-generated images together with photographs.

Walton's often-overlooked second criterion, vulnerability to visual rather than verbal errors, also groups seeing in a photograph together with direct seeing. But it groups seeing a painting and drawing together with those, and it distinguishes AI-generated images as vulnerable to both kinds of error.

Without committing on the question of whether photographs are *fully* transparent, the second criterion allows us to say that AI-generated images are not as transparent as photographs are.

# References


[1] Noël Carroll. Towards an ontology of the moving image. In Cynthia A. Freeland and Thomas E. Wartenberg, editors, *Philosophy and Film*, pages 68–85. Routledge, London and New York, 1995.

[2] Dan Cavedon-Taylor. Photographically based knowledge. *Episteme*, 10(3):283–297, September 2013.

[3] Jonathan Cohen and Aaron Meskin. On the epistemic value of photographs. *Journal of Aesthetics and Art Criticism*, 62(2):197–210, Spring 2004.

[4] Gregory Currie. *Image and mind: Film, philosophy and cognitive science*. Cambridge University Press, 1995.

[5] Cara A. Finnegan. Recognizing Lincoln: Image vernaculars in nineteenth-century visual culture. *Rhetoric & Public Affairs*, 8(1):31–58, Spring 2005.

[6] Oliver Wendell Holmes. The stereoscope and the stereograph. *The Atlantic Monthly*, pages 738–748, June 1859.

[7] William James. The will to believe. In Alburey Castell, editor, *Essays in Pragmatism*, pages 88–109. Hafner Publishing Co., New York, 1948. Originally published June, 1896.

[8] Christy Mag Uidhir. Photographic art: An ontology fit to print. *Journal of Aesthetics and Art Criticism*, 70(1):31–41, Winter 2012. doi:10.1111/j.1540-6245.2011.01496.x.

[9] Aaron Meskin and Jonathan Cohen. Photographs as evidence. In Scott Walden, editor, *Photography and Philosophy: Essays on the Pencil of Nature*. Blackwell, New York, 2008.

[10] Vivian Mizrahi. Seeing through photographs: Photography as a transparent visual medium. *Journal of Aesthetics and Art Criticism*, 79(1):52–63, Winter 2021. doi:10.1093/jaac/kpaa009.

[11] Bence Nanay. Transparency and sensorimotor contingencies: Do we see through photographs? *Pacific Philosophical Quarterly*, 91(4):463–480, 2010.

[12] Regina Rini. Deepfakes and the epistemic backstop. *Philosophers' Imprint*, 20(24):1–16, August 2020.

[13] Ida M. Tarbell. *The Life of Abraham Lincoln*, volume 2. The Doubleday & McClure Co., New York, 1900.

[14] Kendall L. Walton. Transparent pictures: On the nature of photographic realism. *Critical Inquiry*, 11(2):246–277, December 1984.

[15] Kendall L. Walton. Fotografische Bilder. In Julian Nida-Rümelin and Jakob Steinbrenner, editors, *Fotografie zwischen Dokmentation und Inszenierung*, pages 11–28. Hatje Cantz Verlag, 2012. Translated as "Photographic Pictures", available from the author's website. Accessed October 26, 2023. URL: https://sites.lsa.umich.edu/kendallwalton/wp-content/uploads/sites/225/2015/01/walton_photographicpictures.pdf.