On the Expressive Limitations of Prompt-Directed Image Generation

Musa Haydar
Independent Researcher
musa@musah.net

Abstract: Though images produced with generative AI today are often met with negative reception, Kathryn Wojtkiewicz has convincingly argued that these images may still be considered works of art. To do so, she relies on Claire Anscomb's notion of creative agency pertaining to automatic processes and leverages an analogy between generative AI and photography. Wojtkiewicz's additionally argues that the causes for distaste towards generative AI are social, and not aesthetic, ones. Her analogy between photography and generative AI relies on the fact that the artist can predict the outcomes of the process to realize creative agency. However, in her analogy, she hand-waves the differences between these processes, differences which affect how well a generative AI user can predict the resulting features of the image. In particular, photography begins with something seen before the camera, whereas generative AI begins with a linguistic prompt, introducing ambiguity about the user's intentions. Furthermore, the user iteratively adjusts the prompts and selects the image's features from what was generated, and thereby expressing fewer of the user's beliefs, attitudes, and emotions than other mediums. Consequently, we find that these limit the potential for representational expression in a work created with generative AI as compared to other artistic processes.
Keywords: Artificial Intelligence, Art Creation, Expressive Limitations, Image Generation

1. INTRODUCTION

Software capable of generating images based on written prompts has emerged during the past few years, including Dall-E, Midjourney, and Stable Diffusion. This has raised many questions about the potential roles of generative artificial intelligence in the processes of art in both popular and philosophical discourse. These tools are referred to as generative AI, as they involve deep-learning models trained on massive datasets to produce new images that best match the provided text prompt. This paper is specifically about prompt-directed image generation using these tools; other image generation techniques, as well as other forms of so-called generative AI (such as large language models or models trained on small,

focused datasets), are outside its scope.[1]

Although it remains a contentious issue, examples have demonstrated the potential for the artistic appreciation of AI-generated images. One example is *Théâtre D'opéra Spatial*, which was awarded first prize in the 2022 Colorado State Fair photomanipulation competition.[2] Kathryn Wojtkiewicz (Wojtkiewicz, 2023) opens her paper with this example to demonstrate the "frosty" reception to AI-generated images: "their reasoning may differ, but as the news went viral more and more of those who saw it came to the same conclusion: there is something wrong with calling an AI-generated image a work of art." Wojtkiewicz, however, disagrees: she argues that images produced with generative AI may qualify as works of art by relying on Claire Anscomb's (Anscomb, 2021) definition of "creative agency" as it pertains to images generated via "automatic processes." Furthermore, Wojtkiewicz leverages the commentary made about *Théâtre D'opéra Spatial* in public discourse to argue that the negative reception is due to the social backdrop which presently surrounds generative AI, and not due to the artistic qualities of the images themselves.

Although her work matters towards understanding this negative reception, she neglects to consider if there are also aesthetic reasons that might contribute. In particular, Wojtkiewicz's work depends on an analogy between photography and generative AI as automatic processes. Without further discussion, her analogy can be taken as ascribing both processes the same capabilities of artistic expression. Her argument relies on the idea that, in prompt-based image generation as in photography, creative agency depends on a person's ability to predict the results of her actions on the features of the work. However, this is the point at which the processes differ, and in this paper we shall explore the process of prompt-directed image generation to investigate the possibilities of expression in the creation of these images. In particular, we shall focus on the species of artistic expression through visual representation of the subject matter, which we shall refer to as representational expression. When audiences take interest in representational art forms, they are interested in how the artist represents their subject matter, and what that communicates about the artist's beliefs and attitudes towards the subject.

We find two features of prompt-directed image generation that limit the potential for representational expression in the resulting image. First, unlike

---

[1] It is assumed throughout this paper that generative AI plays the role of tool of image creation, whereas it has been argued that it may be understood rather as a machine-collaborator role. We shall return to this assumption and its consequences in the conclusion.

[2] See Kate Knibbs, "Why This Award-Winning Piece of Ai Art Can't Be Copyrighted."

Anscomb's examples of automatic processes which begin with a real object, use of generative AI tools begins with a linguistic prompt, which introduces linguistic ambiguity about what the user intends to capture in the image-making process. Secondly, the users of generative AI do not themselves create the representational features of the image, but instead select what is closest to their expressive goals from the results of the linguistically ambiguous prompt. These two features restrict the user's ability to reliably predict the outcome of their actions, and thus, images produced by generative AI express fewer of the users' beliefs and attitudes about the subject of their images than other image-making processes, including both traditional media (drawing, painting) and automatic processes (such as photography).

In §2, we present Wojtkiewicz's argument supporting the possibility for images produced with generative AI to be considered works of art, including a discussion of Anscomb's (Anscomb, 2021) work on creative agency in automatic processes, upon which it relies. We also rehearse Wojtkiewicz's points about the negative reception being due to the social backdrop of generative AI, and her analogy between photography and generative AI, noting where there is room for further discussion. In §3, we briefly examine theories of artistic value, so as to lay a foundation upon which we may consider how the representational quality of art is evaluated. This in mind, in §4, we continue investigating the analogy between photography and generative AI, focused on a recent paper by P.D. Magnus (Magnus, 2023); this paper has many issues, which are addressed by the work of Wojtkiewicz and Anscomb, but understanding what Magnus missed points us towards understanding expression in generative AI.

In §5, we elaborate on the process of prompt-directed image generation. We find two features in particular that differentiate it from other artistic processes: linguistic ambiguity introduced by the prompt, and selection among the generated features. We conclude that since the final image presents neither the prompt nor the process by which the details were selected in the image, the viewer cannot discern *how* the user chose to represent the subject in the final image. This is the major contribution of this paper. This being how Anscomb explains the aesthetic interest taken in a representational medium (discussed in §3), we understand why the images produced by this process of prompt-based image generation may be evaluated as lacking in representational expression. This is not to discount the possibility of artistic creation that embraces these limitations. However, we argue that when audiences evaluate images for their quality of representational expression, these features in the process of prompt-

directed image generation may be a cause for the distaste towards AI-generated images.

2. CAN GENERATIVE AI MAKE ART? THE CASE FOR AUTOMATIC IMAGES

The question of whether images generated using prompt-driven deep-learning models could possibly be considered art was answered convincingly by Kathryn Wojtkiewicz (Wojtkiewicz, 2023). Wojtkiewicz argues that if the user of a generative AI tool has artistic intent and takes actions to fulfill this intent which demonstrate creative agency, the work qualifies to be considered art. To support the second condition, she relies on Claire Anscomb's (Anscomb, 2021) notion of "creative agency" in the creation of "automatic images." Automatic images, for Anscomb, are those produced via processes that "reduce or replace, to varying degrees, the labor that is required of the hand, mind, or eye of a human artist." Her examples of such processes vary, and include nature printing, camera obscura tracing, and, of course, photography. Because these are "mind-independent" processes, wherein the features of real objects are registered in the resulting work, they are automatic processes.

Anscomb provides a historical overview of the arguments for and against considering works that incorporate automatic processes to be artworks, but eventually comes to this point: when we consider an object a work of art, the necessary condition is that some agent(s) expressed "creative agency" in its making. The automatic processes listed above, for example, are all used by human artists to realize some artistic intent when incorporated into art-making. In Anscomb's words, "the agent works with broad intentional goals toward which they take an executive attitude" and takes actions where they can anticipate what kinds of "formal and representational features the image will have, and what kind of processes will be most appropriate to realize these." These actions may directly modify the art object, as in sketching or sculpting, but even if the art object is created by a remote process, such that the features of the work are not directly created by the agent, it may still qualify as an execution of their intentional control.

An example Anscomb provided of remote actions that realize intentional control is Peter Paul Rubens's delegation of the painting process to apprentices. Though he was no longer directly modifying the painting, Rubens's apprentices acted based on his instructions, thus realizing his artistic intentions in the work. Rubens's paintings are therefore expressive

of his artistic intentions.[3] Likewise, the use of automatic processes may involve remote actions on the part of the artist. Anscomb further points out that the very choice of automatic processes by the artist, and their choices of how to implement such processes, affects the resulting features in the work, thus exhibiting their intentional control.

Anscomb proceeds to show that when viewers take an aesthetic interest in the image as a work of art, they are interested in how the object was represented by the artist, not merely by the object which is represented in the image. Consider Andy Warhol's photo-screenprints, such as his images of Elvis, which used half-tone photographs and large meshes to create the "slightly degraded look of mechanically reproduced cheaply printed images in the mass media" (Potts, 2014, cited in Anscomb, 2021). Anscomb points out that Warhol could have chosen finer meshes to produce higher quality images, or even coarser meshes to the opposite effect, and these possibilities are considered when viewers evaluate the intentional control exercised in the work by Warhol. We shall return to this idea in the following section, but this grounds the fact that automatic images may be, and indeed have been, considered works of art.

Wojtkiewicz builds upon Anscomb's framework, arguing that the use of generative AI is a new kind of automatic image-making process: "The only movement of a human hand required is that of typing the prompt and clicking 'go'." There is an important difference between generative AI tools and the various automatic processes explored by Anscomb: generative AI does not depend on a "real object" whose features are captured by the process. It is instead an automatic process by which the user's intentional control is realized through a linguistic prompt, and the result relies on generative AI model and its training dataset. Wojtkiewicz admits that the analogy between generative AI and Anscomb's automatic processes is "not a perfect one" because of this difference. She defends her positing generative AI as an automatic process on the ground that the dependance, whether on a real object or on a dataset comprising many existing images, is similar, and, furthermore, that its use is a remote action by which one can realize intentional control.

Indeed, despite their misnomer as tools of "artificial intelligence," the image generation tools in question provide no intentionality of their own. Rather, they realize a human intention in generating an image described to

---

[3] This is not to say the work exclusively expresses Rubens's intent. Indeed, various degrees of Rubens's apprentices' intentional control may also have been expressed in the final work.

them by the prompt.[4] Wojtkiewicz explains, then, that "if a sufficiently skilled user of a generative AI program can anticipate the effect their prompt will have on the image created, then they have intentional control over the outcome just as a photographer does when they adjust the settings of their camera." In particular, the generative AI user is able to iterate upon and adjust their prompt in anticipation of the expected result. To summarize: since the images are produced by a user with an artistic outcome in mind, a strategy to achieve the outcome using the processes of their choice, and undertakes the techniques with intention to create the features of the desired outcome, then it must be "at least possible for AI-generated images to be artworks" (Wojtkiewicz, 2023).

Presently, images created with generative AI tools are typically received with controversy. Wojtkiewicz claims the reason for the distaste for generative AI as a social one and not an artistic one. She elaborates three specific concerns about generative AI that account for this distaste: first, that the training on massive image-data sets without the consent of the artists may amount to plagiarism; secondly, that considering generated images as artwork may "[cheapen] more traditional mediums," as viewers may come to undervalue the skill required to produce such images manually; and thirdly, that generative AI may cause the "overall decline of other artistic mediums" as individuals may prefer the ease of generating an image over the effort and skill needed to produce an image in traditional mediums. We will not rehash her discussion here, although these are all worthy points towards understanding how generative AI has been received. Indeed, the social backdrop of generative AI is quite contentious.[5]

Wojtkiewicz's argument is sound insofar as countering the exclusion of generated images from consideration as art altogether. However, especially on account of her analogy between photography and generative AI, one could read her paper and come away with the idea that images produced with photography and with prompt-directed image generation are similarly capable of artistic expression. Wojtkiewicz's hand-waving the difference between the prompt-based image generation and Anscomb's examples of automatic process, which depend on and capture the features of "real objects," is suspect here.

---

[4] The data used in training an AI model may also contribute to the intentional control which determines the properties of the resulting image. See Anscomb, "Creating Art with AI," for examples and further discussion on this point.

[5] One might add to Wojtkiewicz's list that another glaring cause for the distaste is due to the unsettling quality of the images generated by the technology in its present state. This is yet another social reason rather than an aesthetic one, and is not taken up in this paper. The point of whether audiences think images produced with generative AI are aesthetically pleasing is independent of the process by which a person realizes artistic intent.

Her analogy between generative AI and photography is as follows: in both processes, the artist first chooses the medium and tools which she believes will allow her to express her artistic intent. Then, she uses her knowledge of the medium to determine which techniques allow her to best achieve this result (for photography, adjusting the "aperture, lighting, lens type, and so on"; for generative AI, this might be various prompting techniques). A key point in Wojtkiewicz's defense of AI-generated images is that the user is able to "reliably predict" the effect the prompt has on the image. If the desired features fail to be created by these tools and techniques, the artist may then repeat the process until her intention is achieved, and through all this, she is exhibiting creative agency. "All this suggests the final product exists because of the [artist's] intentional action. Thus, the resulting [image] passes our threshold for an artwork."

This being true, we might be tempted to conclude that, since both processes involve remote realization of the artist's intent in this iterative manner, that the processes are similarly expressive of that intent. The point we would be neglecting if we made such a conclusion, however, is that degree to which the artist has intentional control over the features of the work, respective to each process. Here, we will argue that an artist has more or less intentional control over the work depending on her ability to predict the outcomes of the process. When taking a photograph, for instance, a photographer can see directly before her what she will capture with the camera. Beginning with what one can see allows her to predict, with a high degree of reliability, how the features of the object will be captured in the work, with respect to her skills, the photographic technology she has chosen, and so on, even if she cannot entirely predict the resulting qualities of the photograph.

This is not the case for generative AI, where the user begins with a written prompt. Anscomb herself had previously written that "AI agents can generate formal properties that artists may not fully anticipate in advance of their creation." (Anscomb, 2022). This is important when evaluating the users ability to predict the outcomes of their actions, and therefore the degree to which they can realize intentional control. The point here is that the unpredictability detracts from the artist's ability to realize their intent, at least insofar as their intent involves representation. A discussion of expression using prompt-based image generation processes requires, firstly, a formalization of what we mean by representational expression, which will be the next section. Following that, we shall more carefully consider the process undergone by the artist when prompting and generating an image using generative AI and, to carry on with this analogy,

compare that to the process of composing a photograph.

3. BRIEFLY, ON EXPRESSION AND VALUE IN ART

There are, of course, various categories of what an artist may express in a work as well as what features an audience may look to for evaluating it as a work of art. Anscomb (Anscomb, 2022) presents two conceptions of artistic value: the monist conception, where artistic value derives solely from aesthetic value, and the pluralist conception, as explained by Robert Stecker: "artistic value is a function of, and derived from, a plurality of more basic values, including, but not confined to, aesthetic value (Stecker, 2019). Artworks are also valued as artworks for their cognitive value, ethical value, art-historical value, interpretation-centred value, and in other ways as well". Naturally, the monist conception is quite narrow as it neglects a whole gamut of conceptual art history and works appreciated for their cognitive value over their aesthetic value (if the latter is present at all). Any number of these artistic values may be the the artist's intention in a given work; likewise, audiences may evaluate and appreciate a work for any number of these values, regardless of the artist's intentions. For instance, one might find the images produced by Joseph Kosuth for his 1966 conceptual piece, *Titled (Art as Art as Idea),* to be aesthetically beautiful, despite simply consisting of enlarged photographs of linguistic definitions such that the work, according to Kosuth, is the concept and not the images themselves.

The question of artistic causes for the distaste towards AI-generated images, with which this paper is concerned, will pertain to representational expression. This refers to an artist's representation of an object or idea in the medium. Robert Hopkins (Hopkins, 2015) presents a formulation of "communicative representational art" and explains it as follows: "Where there is representational art, there had better be representation, and its presence had better make a difference to what interests us. Where representational art is communicative, what interests us (and what representation furthers) is the communication of thoughts." Hopkins's restriction to the communication of "thoughts" is debated by Paloma Atencia-Linares (Linares, 2018), who instead concludes that it is not necessary to communicate thoughts but rather "to make the intentions manifest," a similar notion to Anscomb's creative agency.

On the other hand, audiences are not merely interested in the presence of creative agency, but in what the intentions and agency of artists presents

about them and their represented subjects. Anscomb explains that when viewers take an "aesthetic interest" in representational art forms, it "is not merely to take an interest in the object that is represented, but *how* it is represented as the result of an artist's intentional activity." This interest includes evaluations of what the representational work presents of the artist's attitudes, beliefs, or emotions about the subject of her work. Even for automatic images, "through the activity of the agents responsible for the production of these images, we are shown the objects as they are conceptualized or seen by the agents in question" (Anscomb, 2021).

We will show in the following sections that the processes chosen by the artist will affect how the art object expresses their beliefs and attitudes, and that prompt-directed image generation expresses fewer of these beliefs than other mediums. This is not to say that any limitations of representational expression in AI-generated images are a problem *per se*. A work of art does not necessarily need to be representational, something much abstract and conceptual art has explored.[6] As Sol LeWitt put it, for conceptual art, "what the work of art looks like isn't too important" (LeWitt, 1967). This is to say, conceptual expression, among other artistic values, is not limited as we find the representational expression in AI-generated images is: conceptual art often leverages just such unpredictability and ambiguity. There is a profound history of work by conceptual artists' experimentation with whether the object of art can be "nominal, linguistic, invisible," and so on (Dworkin & Goldsmith, 2011). AI-generated images may likewise foreground artistic values beyond aesthetic or representational categories.

The questions of what is possible for conceptual art with AI-generated images, or what other artistic values may be foregrounded by generative AI, will not be considered further in this paper (and will more likely be discovered by artists than theorists). Rather, we argue that insofar as images are evaluated for what they express through visual representation, as Anscomb argues images created by both traditional and automatic processes are, the limitations involved in the process of generating images via linguistic prompts may be a cause for the distaste towards AI-generated images.

4. COMPARING PHOTOGRAPHY AND GENERATIVE AI

Much of the recent literature on this topic draws an analogy between

---

[6] See Dworkin, "The Fate of Echo," on conceptual art.

photography and generative AI, as we have already seen done by Wojtkiewicz. There are a few reasons for this: first, both are automatic processes which create an image, possibly requiring quite little effort. Secondly, during the early days of photography, debate arose over whether photographs could be considered works of art at all. Cynicism towards photography by art critics and theorists is often cited as towards a defense of generative AI.[7] Towards understanding how the processes differ in their potential for representational expression, we want to investigate what degree of intentional control is possible for each process, respectively.

Along the same line as Wojtkiewicz, P.D. Magnus likens the use of generative AI to photography by considering both processes to be "mechanical" (Magnus, 2023). This notion relies on Kendall Walton's explanation that, since "photographs depend counterfactually on the objects photographed, even when the beliefs and other mental states of the photographer (and anyone else) are held fixed," they are produced by a mechanical process (Walton, 2012, cited in Magnus, 2023). For generative AI, Magnus writes that "I entered a prompt, and the algorithm did the rest. So the image is mechanical". This notion is essentially equivalent to Anscomb's notion of automatic processes, at least as far as this construction is concerned, and we will prefer Ansomb's terminology.

The point of Magnus's paper is to address the question of whether images produced with generative AI are transparent: according to Walton, transparency is the quality of a photograph such that, when looking at a photograph, it is as though one is peering through the photograph onto the original object. Magnus concludes that generative AI images lack transparency: the process does not record an object directly, and image generation relies on a verbal description which is prone to errors. While this conclusion is agreeable (in fact, it seems quite obvious), the question of transparency is not one we're concerned with; yet, the problems with Magnus's discussion of photography and generative AI shall give us some direction.

To begin with, Magnus's claim that "what is included in the photograph does not depend on the photographer's beliefs about what is in front of the camera" is incorrect. Magnus would be mistaken to have derived this from Walton, and the theory that photography is belief-independent has been well argued against. For instance, Dominic Lopes (Lopes, 2015) puts

---

[7] As examples, see Blaise Agüera y Arcas, "Art in the age of machine intelligence" and Arthur Still and Mark d'Inverno, "Can machines be artists? A Deweyan response in theory and practice."Anscomb likewise points out the frequent relation of photography and AI, stating "I think however, that a different parallel is more fitting: there is a historical tendency to overplay the reaches of new imaging technologies" (Anscomb, 2022).

forward an argument that "if photographs are images that depict by belief-independent feature tracking, then they are no different in this respect from drawings made by belief-independent feature-tracking," establishing similarity between the two processes to demonstrate how each may realize (or fail to realize) creative agency.

As we discussed in §2, Anscomb's own formulation of creative agency demonstrates how the artist's intentional control is still present in photographic work.[8] For instance, the photographer's decision to compose the photograph at all may betray a belief that the thing is worth photographing. Then, the photographer's role in arranging the scene, adjusting the lighting or the lens aperture, framing of the subject, and so on, similarly indicate some beliefs or attitudes about the subject. Each potential adjustment (and revision, if the initial photographs were unsatisfactory) is some means by which the photographer may express, through the image, her beliefs and attitudes about the object of her photograph. Even the choices of which photographic technologies to employ may reflect some beliefs about the subject matter, insofar as, according to Anscomb, it exercises intentional control over the features of the image.

Magnus makes the same incorrect claim about images produced with generative AI; here, Wojtkiewicz has addressed the extension of Anscomb's work to generative AI. In particular, intentional control is realized in producing an image with generative AI on account expression in the prompt and the selection of the final image. A running example in Magnus's paper is of generating an image of Abraham Lincoln wearing "a funny hat" (his prompt in full was "Abraham Lincoln in a funny hat; portrait; historically accurate"). He writes that, in the first image generated, the hat appears "disappointing, [...] like nothing more than a pom pom." That he was disappointed at all indicates that Magnus holds some idea of what represents "a funny hat," that he sought to express this idea of a "funny hat" through his prompt, and failed to do so in his first attempt. Another user would therefore be likely to adjust her prompt, perhaps specifying what she believes to be a "funny hat" in more detail ("Abraham Lincoln wearing a helicopter hat; portrait; historically accurate"). Then, the images, still generated by the same automatic process, would be nearer to

---

[8] Lopes (Lopes, 2015) points out the counterargument that photography (and similarly for generative AI) is belief-independent because it captures features beyond the photographer's intention. For instance, one photographing the Grand Canyon did "not intend to depict the speck of a vulture hovering above the canyon floor." He shows that this counterargument too strongly binds "agency" to "intention." Here, Anscomb's point is that creative agency is exhibited through intentional control, and not that the outcome is the sole or complete realization of the intention in the work.

her beliefs about what makes a hat funny, or what hat would be funny for Lincoln to be portrayed wearing.

Magnus acknowledges this objection: he describes the process by which one revises the prompt to generate a new output, "which can either be accepted as good enough or the basis for further runs." For the sake of his argument on the transparency of generative AI, he proceeds to dismiss this, saying the example he provided was the "first result generated by the algorithm given my initial prompt." His reasoning is that, in doing so, most of the work had been done by the generative AI system. However, his deliberate choice not to revise his prompt cannot negate his creative agency in the process, especially because the initial image was generated based on a linguistic prompt of his making. Magnus himself admits that, under such an iterative process, "the final image ultimately relies on [one's] beliefs and attitudes," but this neglects his role in generating the initial image. Indeed, he had something in mind (ideas about Lincoln and funny hats) which he firstly chose to produce an image of with generative AI and secondly took an action to describe it. Even a single prompt demonstrates intentional control when considering alternative words or phrases Magnus might have chosen (for instance, a "silly hat" or "weird hat" instead of a "funny hat" may correspond to varying intentions). He further claims that, in being an iterative process, the process differs from the way a "single photograph" is produced; however, we have already shown that even the production of a single photograph demonstrates an artist's creative agency.

Thus, we have explored, for both photography and generative AI, how an artist may express creative agency. In producing a photograph, the artist begins with something before his eyes, the object being photographed, and exercises agency in determining how the features of the image will appear in the photograph. On the other hand, when using generative AI, creative agency is expressed in writing the prompt and selecting the image, or for particular features in the image. This is the same reason Magnus found that photographic transparency does not occur in generative AI: an image generated with these prompt-driven tools begins not with a real object but instead with a linguistic description. In the next section, we shall consider further what these actions might and might not express about the subject matter represented in the work.

5. THE PROCESS OF PROMPT-DIRECTED IMAGE GENERATION

Let us once again elaborate the process of prompt-direct image

generation: suppose one has in mind something she wishes to express via an image—after all, how could one compose a prompt to describe an image without having something she wishes to express in mind? Having chosen from among the available generative AI tools, she writes a prompt which describes the image to whatever level of detail she expects will produce the image she desires. To return to Magnus's example, perhaps she writes "Abraham Lincoln in a funny hat; portrait; historically accurate." The tool then outputs a few images which it has determined, by its model and training data, fits the description provided: various portraits of Lincoln, to various degrees of historical accuracy, wearing various hats. The user may then select one or more of these images, to put to whatever use she intended.

If none of these images are satisfactory to her end, she may adjust her prompt and generate more images. For instance, perhaps she desires a colored image, whereas all the images generated were monochrome (as most images of Lincoln are), so she revises the prompt to be "Abraham Lincoln in a funny hat; colorized portrait; historically accurate." With each round of images generated, the user perceives the changes in the resulting images, and adjusts or reforms her prompt until the results are eventually satisfying to her end (or perhaps they never are, and she finds generative AI to be inadequate, and perhaps lacking expressivity). Again, it is because Magnus considered only the object produced—the photograph or the generated image—and neglected the intentional control of the artist, that his accounts fail when considering expression in both processes.

Notably distinct in this process is that the features in the resulting image were not created by the artist directly. A painter, for example, directly modifies the painting with each stroke in the process until the painting satisfies his artistic intentions. When using these prompt-driven generative AI systems however, it is not the image that is being formed by the user, but the prompt. While the subject matter of the image is described via the prompt, and while the user iterates on it to eventually produce the final image, the particular qualities of the image—the details included or omitted, the colors and textures, etc.—are not created by this process. Instead, the user selects the most adequate image from the options generated.

From the above account, we have observed two things that distinguish the use of Generative AI as an artistic process. The first is the presence of transmedia interaction: the user composes and iterates upon a linguistic prompt in order to describe the features of the image to be generated. The second is that the user does not create the features of the image directly,

instead iteratively modifying the prompt until the images generated are as close as possible to what he would like to express through visual representation. We shall show that, as a result, the image cannot represent the beliefs and attitudes about the subject that the user would express to the same degree as other mediums.

5.1 Linguistic Ambiguity in Prompts

Relying on language to describe the image to be produced introduces inevitable ambiguity about the user's representational intentions. When language is itself the medium of art (as in poetry or conceptual writing), the ambiguity is taken as a feature of the work. This is true of Lawrence Weiner's 1968 exhibition, *Statements*, which "contained two dozen self-descriptive pieces composed of short phrases, grammatically suspended by the past participle without agent or imperative, as if they had already been realized as soon as written (or read)" (Dworkin & Goldsmith, 2011). These statements, Dworkin explains, were themselves the works of art. The ambiguity in a statement such as "one aerosol can of enamel sprayed to conclusion directly upon the floor" is a feature of the work. Weiner declared his intentions in his "Declaration of Intent" and among them is that "the work need not be built." Although one may try to interpret or imagine these statements, that is not their point *per se*.

Contrast this to a prompt composed for generative AI. In the image generation process discussed in this paper, the user's intention is that the language describes something which will be represented visually. However, in the end the art object is not the user's prompt, but the resulting generated image. If one was to generate an image of "one aerosol can of enamel sprayed to conclusion directly upon the floor," then the matter of what the can looks like, what the enamel looks like, what material the floor is, how the enamel interacts with it, and so on, are all ambiguous. So, if one were to begin with such a prompt, she may then have to adjust it for each parameter if the result is not satisfactory, but there is simply no means to eliminate all the ambiguity, and some details, or at least how these details are represented, will be left to chance. This chance here is, of course, determined by the generative AI model and the dataset with which it was trained. For large models such as Dall-E and Midjourney, these datasets are not available and are incomprehensibly large, such that even if this data were listed, it is unlikely that one could review them to better predict what results the model would produce.

One might point out that as a user's familiarity increases with the model, her ability to predict how the model will represent certain features will

improve. For instance, many have become familiar with generative models' difficulty accurately representing hands, and therefore prompt for images of people where their hands are not shown.[9] However, this is true of all artistic mediums: the more skilled a photographer is, the more she is capable of predicting how her camera, lighting, scene, etc. will affect the features of the photograph. Likewise with traditional mediums: the more skilled the painter, the more reliably she can predict how her choice of paints and brush strokes will affect the resulting painting. Despite the potential for improvement in predicting the outcomes, we have demonstrated that linguistic ambiguity is an inevitability for prompt-directed image generation. This ambiguity necessarily creates a degree of unpredictability in the outcomes of the user's actions, as compared with photography, painting, and many other mediums.

5.2 The Process of Selection

Furthermore, the artist expresses herself by selecting among things she did not create herself. Instead, she tries to get as close as possible to the ideas she wishes to represent in the work. Had she created the details in her work by directly shaping the material, one could argue that an artist would still be unable to exactly express the idea she has about her subject, and seek to get as close as possible to a satisfying result. Indeed this is true especially due to the limitations of her skill or of her chosen medium. The point here is not that these processes precisely or entirely express her beliefs and attitudes in the work. Rather, the process by which one works creatively with the material and medium embeds her ideas in the work, both consciously and unconsciously. Consider collaging as an example: when collaging from existing illustrations or photographs the artist may not even be aware of her idea of the style, proportions, or details that make a hat "funny." When selecting an image of hat for her collage, she will choose the hat available to her that comes closest to this. However, if she was asked to draw a funny hat on a photograph of Lincoln by hand, many of these ideas of what makes a hat funny (or specifically funny for Lincoln to wear) would necessarily be expressed.

After having composed a prompt and generated some images, the user of generative AI will select whichever she finds most suitable. It becomes clearer to an observer what she is trying to express when he considers the options that were selected against. In *Art and its Objects*, Wollheim (Wollheim, 1980) explains Gombrich's argument of how an object

---

[9] See Kyle Chakya, "The Uncanny Failure of AI-Generated Hands."

becomes expressive in the context of alternatives:

> a particular element has a significance for us only if it is regarded as a selection out of a specific set of alternatives. Blue has no significance: blue-rather-than-black has: and so has blue-rather-than-red though a different one. In the light of this, the notion of 'context' can be made more specific. In order for us to see a work as expressive, we must know the set of alternatives within which the artist is working.

Wollheim presents this argument to counter the claim that the expressiveness of an object can be determined solely from itself. Consider that every stroke of a painter, for example, could be thought of among the alternative strokes the painter did not take. Of course, when appraising a work of art, we generally do not consider every counterfactual possibility to such a degree. But we often consider, at a higher level, the artist's decisions of inclusions, exclusions, placements, etc. as compared to hypothetical alternatives.

When one did not create the components of the work, but instead selected them, what the user first described in the prompt and then selected against becomes more significant towards evaluating the visual representations in the work. However, neither the prompt used to generate the image, nor the alternatives selected against, nor the iterations performed by the artist, are typically presented with the final AI-generated image. It is difficult to consider the hypothetical alternative decisions made by the artist when one does not know what decisions the artist did make.

6. CONCLUSION

Thus we have demonstrated two particular features of the process of prompt-directed image generation that differentiates it from other processes of art-making. The first is that the user begins the process with an ambiguity-introducing linguistic prompt which describes what she intends to represent visually. The second is that she proceeds to select features created by the model which interpret her prompt. As discussed in §3, Anscomb explains that when viewers take an aesthetic interest in a representational form of art, they are interested in how it was produced by the artist. Unless the viewer is presented with the prompt and perhaps the alternative images which the user selected against when generating her final work, one cannot discern the choices made when generating the image. Furthermore, even if the image is presented with the prompt, the viewer still cannot discern to a high degree what was intended to be visually

represented in the ambiguity of the prompt. This poses a limitation on the representational expression of prompt-directed image generation, and this may very well be an artistic reason for the distaste with which generative AI images are met. It is more than the social backdrop: the limitation is intrinsic to the process by which AI-generated images are created, and it limits the depth to which representational expression is possible in the work.

None of this has been to argue that images produced with generative AI cannot or ought not to be considered works of art. Wojtkiewicz has indeed convinced us that they can. As curation and selection are themselves kinds of expressive acts, audiences may take interest in the process by which an AI-generated image is created. For instance, Jason Allen, creator of the *Théâtre D'opéra Spatial,* claims to have composed 624 text prompts in the creation of the image. Perhaps appreciation for this process of selection could become the grounds of evaluation for AI-generated images.

Furthermore, just as photography gave rise to new categories of artistic evaluation, the increasing use of these deep-learning tools may likewise give rise to new evaluative frameworks. It has been assumed throughout this paper that the role of generative AI is that of a tool. This is commonly taken to be the case since generative AI software exhibits no intentions nor agency of its own. Nevertheless, reframing the role of generative AI to be that of a collaborator (Anscomb, 2022) or participant (Cross, 2024) could allow one to view AI-generated images not as lacking in expression of the user but instead as presenting human intention with machine collaboration. Nick Young and Enrico Terrone (Young & Terrone, 2025), for instance, propose rethinking of the role generative AI plays in image creation: they argue that generative AI is "not an agent which creates, or contributes to the creation of, art, neither, strictly speaking, is it a tool which artists use; rather, it is more similar to a plot of land in which a gardener might sow seeds." Evaluating images produced with generative AI as such, rather than as representational art images created by the user of a tool, might yield more positive outcomes.

While such reframings may be possible, generative AI is often thought of as a tool of image creation and its products are evaluated similarly to other artistic mediums. When an evaluation of an image relies on how the artist creates representations and how these representations express the ideas, attitudes, and emotions of the author, audiences may find AI-generated images lacking. Again, this is not to discount the possibility of appreciation of the cognitive value or even aesthetic beauty of an image,

but it is to say that representational expression being a primary category for evaluation may be the cause for distaste towards these images. This is, of course, in addition to the ethical and social evaluations put forth by Wojtkiewicz. Nonetheless, as time passes and artists continue to experiment with the technology, the very values with which we evaluate works of art may evolve, just as conceptual art brought to the forefront the cognitive value in art. As Walter Benjamin famously examined how the mechanical reproduction of images has "gained a place for itself among artistic modes of production" (Benjamin, 2018), generative AI may likewise find its place in the processes of art.

References

Agüera y Arcas, B. (2017). Art in the age of machine intelligence. In *Arts* (Vol. 6, pp. 18). MDPI.

Anscomb, C. (2021). Creative agency as executive agency: Grounding the artistic significance of automatic images. *The Journal of Aesthetics and Art Criticism*, *79*(4), 415-427.

Anscomb, C. (2022). Creating art with AI. *Odradek. Studies in Philosophy of Literature, Aesthetics, and New Media Theories*, *8*(1), 13-51.

Benjamin, W. (2018). The work of art in the age of mechanical reproduction. In *A museum studies approach to heritage* (pp. 226-243). Routledge.

Chakya, K. (2023). The Uncanny Failure of AI-Generated Hands. *The New Yorker*.

Cross, A. (2024). Tool, collaborator, or participant: AI and artistic agency. *British Journal of Aesthetics*, ayae055.

Dworkin, C., & Goldsmith, K. (2011). *Against expression: an anthology of conceptual writing*. Northwestern University Press.

Hopkins, R. (2015). The real challenge to photography (as communicative representational art). *Journal of the American Philosophical Association*, *1*(2), 329-348.

Knibbs, K. (2023). *"Why This Award-Winning Piece of Ai Art Can't Be Copyrighted."*. Wired.

LeWitt, S. (1967). Paragraphs on conceptual art. *Artforum*, *5*(10), 79-83.

Linares, P. A. (2018). Is photography really limited in its capacity to communicate thought? A response to Hopkin. *Laocoonte. Revista de Estética y Teoría de las Artes*(5), 83-96.

Lopes, D. M. (2015). A New Theory of Photography. *In Four Arts of Photography*, 65–86.

Magnus, P. (2023). Generative AI and photographic transparency. *AI & SOCIETY*, *40*(3), 1607-1612.

Potts, A. (2014). The Image Valued 'As Found'and the Reconfiguring of Mimesis in Post-War Art. *art History*, *37*(4), 784-805.

Stecker, R. (2019). *Intersections of value: Art, nature, and the everyday*. Oxford University Press.

Still, A., & d'Inverno, M. (2019). Can machines be artists? A Deweyan response in theory and practice. Arts,

Walton, K. L. (2012). Fotografische Bilder. In: Julian N-R, Jakob S (eds) Fotografie zwischen Dokmentation und Inszenierung. *Hatje Cantz Verlag*, 11–28.

Wojtkiewicz, K. (2023). How do you solve a problem like DALL-E 2? *The Journal of Aesthetics and Art Criticism*, *81*(4), 454-467.

Wollheim, R. (1980). *Art and its Objects*. Cambridge University Press.

Young, N., & Terrone, E. (2025). Growing the image: Generative AI and the medium of gardening. *The Philosophical Quarterly*, *75*(1), 310-319.