Artificial Aesthetics

A critical guide to AI, media and design

A book by
Lev Manovich & Emanuele Arielli

Preface

Suppose you are a designer, an architect, a photographer, a video maker, a musician, a writer, an artist, or a professional or student in any other creative field. Or perhaps you are a digital creator making content in multiple media.

You may be wondering how AI will affect your professional area in general and your work and career. This book does not aim to predict the future or tell you exactly what will happen. Instead, we want to offer you a set of intellectual tools to help you better navigate any changes that may come along.

These tools come from several different fields: aesthetics, philosophy of art and psychology of art (Emanuele), and media theory, digital culture studies, and data science (Lev). As far as we know, our book is the first to bring together all these different perspectives in thinking about creative AI.

We started the work on the book in summer 2019, exchanging numerous messages, commenting on each other ideas, and sharing drafts of sections. The final book is a result of this process. Although each chapter is written by one author, it reflects the discussions we had over 27 months.

The book is released one chapter at a time on manovich.net, academia.edu, and medium.com.

Lev Manovich and Emanuele Arielli

November 2021

Artificial Aesthetics: A Critical Guide to AI, Media and Design

Proofreading: Julian Sunley.
Design and cover: Lev Manovich.
Chapters titles font: Baskeville.
Quotes font: Open Sans.
Body font: PT Sans.

The Public Fonts (PT Fonts) are a family of free fonts, designed by Alexandra Korolkova with with Olga Umpelova and Vladimir Yefimov (Russia). The fonts are included in macOS and available in Google Fonts. PS Sans was the first font in this family released in 2009.

Lev Manovich is a a Presidential Professor of Computer Science at The Graduate Center, City University of New York (CUNY). He is the author and editor of 14 books including *Cultural Analytics*, *AI Aesthetics*, *Instagram & Contemporary Image,* and *The Language of New Media* described as "the most suggestive and broad-ranging media history since Marshall McLuhan." His publications are available at https://manovich.net.

Emanuele Arielli is Professor of Philosophy and Aesthetics at IUAV University, Departments of Architecture and Arts, Venice, Italy. His research interests include aesthetics, art theory, semiotics, cognitive science, and philosophy of mind. Arielli most recent books are The Aesthetics and Multimodality of Style (with M. Siefkes, 2018) and Idee Virali (with P. Bottazzini, 2018). For full publications list, see https://iuav.academia.edu/EmanueleArielli.

Chapter 1
Emanuele Arielli

"Even an AI could do that"

What is aesthetics? Consider the many aesthetic choices that we make in our everyday life – picking out and matching clothes, liking photos, choosing a hairstyle, makeup, places to visit, objects to purchase, music to listen to, and so on. In all of these examples, aesthetics refers to pleasurable experiences mediated by our senses. The term can also include concepts such as style and aesthetic judgments that assess the value of an artwork, although the nature of the relationship between aesthetics and art has become an object of debate in contemporary times. We also make everyday aesthetic decisions when creating graphs, capturing and editing photos and videos, drawing images, and designing spaces and buildings. Aesthetics covers both natural and human-made objects and experiences.

Since the beginning of the 21st century, computation, data analysis, machine learning, neural networks, and artificial intelligence (AI) - an all-encompassing and catchy label with a shifting definition - have all gradually entered the aesthetic realm. For example, music streaming services such as Spotify, Apple

Chinese landscape paintings produced using Generative Adversarial Networks. Xue, Alice. "End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks." Proceedings of IEEE WACV, 2021, pp. 3863-3871. arXiv, https://arxiv.org/abs/2011.05552.

Music, and Pandora automatically recommend music we may like. Instagram's Explore tab automatically curates photos and videos in a way that is personalized for each user. Automatic one-button photo improvement is a standard feature in all mobile and desktop apps for editing photos. Large online fashion retailers offer automatic suggestions for additional clothing items, and so on.

These systems rely on increasingly sophisticated methods to predict what people might like. Deep neural networks, for example, learn principles of aesthetic quality by directly observing people's aesthetic choices. Earlier predictions of image quality ratings were based on classical compositional rules (such as the rule of thirds, aspect ratio, saturation, and so on) as well as on the programmers' intuition of aesthetic value, which derived from their observation of the most liked photographs.[1] Later, neural networks were progressively used to assign semantic labels ("meanings") and to automatically extract aesthetically relevant features through the analysis of large databases of liked images.[2]

In addition to recommendations and automatic editing, AI is now widely used to generate new synthetic artifacts, including artworks, music, designs, and texts. For instance, in 2016, a deep-learning algorithm was trained to learn Rembrandt's style by analyzing his 346 known paintings. The algorithm was subsequently given the task of generating a brand-new portrait, the result of which looked uncannily like a real Rembrandt. In the same year, researchers at the Sony Computer Science Laboratories in Paris developed a neural network, called DeepBach, that produces choral cantatas in the style of J.S. Bach.[3] Since then, other music-generating algorithms have been created. Even YouTube videos invite viewers to participate in musical "Turing tests," challenging them to distinguish AI-penned compositions from human ones. For people with some musical training, the task still seems straightforward, but for inexperienced listeners, this is not always the case.[4] In 2019, Deutsche Telekom put together a team of international experts in music and AI to complete Beethoven's unfinished 10th symphony, thus celebrating the 250th anniversary of his birth. The completed symphony, "Beethoven X - The AI Project," premiered on October 9, 2021 in Bonn. It can be challenging to

keep pace with such quick progress as incremental technological changes yield continually improved results: in 2019, an AI used the computing power of a new smartphone model to finish Schubert's "Unfinished Symphony" (n. 8, 1822),[5] although this was accomplished with the help of a composer who cherry-picked the best generated melodies. In 2020, an undergraduate student at Princeton University used a so-called Generative Adversarial Network (GAN) to produce traditional Chinese landscape paintings that were capable of fooling humans in a visual Turing test.[6]

The encounter between AI and aesthetics is crucial because aesthetics is considered a quintessentially human domain. Its intractability and complexity have long appeared as insusceptible to algorithmic reduction. For some, art, aesthetics, and creativity are the pinnacle of human abilities and therefore represent a final bulwark against the seemingly unstoppable advances of AI. In other words, this complex field becomes the ultimate testing ground for AI's possibilities and limitations.

Still, a prevailing opinion holds that developments like those mentioned above just *mimic* existing styles and are not creative at all. In those instances, computers receive pre-existing examples and generate variants conforming to their patterns, while trying to introduce some level of variation. Sometimes they are *uncannily* similar to genuine artworks, but this can also mean that they seem a bit off to a trained eye, lacking the final touches that would make them convincingly human. These algorithms do not generate styles of music or painting that are entirely new, instead they are instances of what we might call *computational mannerism*.

However, it could be just a matter of time until even the experts are deceived and an AI produces artworks that are judged as aesthetically superior to their human variants. One should bear in mind that the examples mentioned above involve artwork sets with a good amount of repetition and low variability: qualities that enable neural networks to extract general features and generate new examples easily. In other words, it seems particularly straightforward to produce traditional or classical artworks as they tend to display a clear, recognizable style and follow the specific patterns of an artist, school,

_____

The encounter between AI and aesthetics is crucial because aesthetics is considered a quintessentially human domain

or tradition. Machine learning systems are ideally suited to analyze numerous occurrences of an object type with slight variations and extract the relevant features and patterns. It would, on the contrary, be very difficult to reproduce something like a Duchamp-style body of work, since the AI would have to start with the very heterogeneous dataset of this artist's oeuvre, encompassing *Fountain*, *Bottle Rack*, the *Large Glass*, the late *Étant donnés*, and so on. Typically, conservative views on art consider technical mastery as a criterion for "real art," and many people still don't consider something that doesn't require technical ability to be art. However, technical ability means procedural knowledge, and AI are designed to deal with precisely this kind of knowledge. Clearly recognizable styles are *well-defined* problems that can be reduced to computational tasks, while the generation of variants that don't follow compositional rules (like Duchamp's works) results in ill-defined tasks that have no easy procedural solution.[7] **"My kid could have done that!"**, the popular cliché directed at contemporary art, seems now, in an ironic reversal, to turn against the great and stylistically complex - but computationally scalable - art of cultural tradition: **even an AI could do that.** It is the Duchamp that remains outside of AI's creative abilities, at least for now.

Here is a brief overview of the main issues that we would like to deal with.

An investigation of the impact of AI and machine learning on aesthetics requires, at the outset, a general mapping of the areas where aesthetics and computational methods meet and relate to one another (see next section, "A simple map"). Then, further on, we will show some points of contact between so-called experimental aesthetics and computational applications, showing

how some limits and critical points found in the former can be transferred to approaches undertaken by the latter (section "Computation and psychology").

Technology is the development of *tools* extending our reach and power. We have biologically limited physical strength: thanks to levers, gears, and eventually engines, we managed to overcome these limits. We have biologically limited visual acuity, but microscopes and telescopes allowed us to amplify the realm of the visible. Similarly, our cognitive skills such as calculation and memory have upper limits, but calculators and computers augmented those skills. Following this line of argumentation, one could suggest that aesthetic capacity has human limits as well, that there could be a point at which peak creativity, or peak aesthetic sensibility, is reached. Limits would be determined by both the individual, who has their own supply of sensitivity, creativity, and skills, and by the culture as a whole, which delimits what is possible within a specific artistic medium. **Artificial aesthetics can be described as an augmentation of our aesthetic skills**, deepening both our creative processes and our understanding and sensibility of cultural artifacts. Advanced systems would then be a further evolution of devices that are already used in creative disciplines, such as graphic programs, computer-aided design technology, music software, and so on (see later chapters on creativity, media theory, and digital culture). If in a traditional sense media are extensions of human senses, then AI is a further extension of human capabilities in mediating between us and the world.

Our engagement with technology expands and modifies how we create and ultimately shapes our cultural evolution. The question arises as to whether all this has the potential to push the boundaries of our knowledge about human cultural and artistic heritage. In a futuristic scenario, machines could acquire a precise understanding of human aesthetic preferences, eventually registering how we perceive and react in front of an aesthetic object with greater accuracy than is available to humans. Machines could learn to produce aesthetic artifacts and generate new creative styles and genres. By analyzing human aesthetics and the diversity of aesthetics in human culture, they may even be able to create new "cultures" - that is, to create genuinely new types of art and aesthetics.

In discussions around AI, we often hear how machines "solve" domains that we thought were uniquely human or achieve better performances than their human competitors. On each occasion, the bar of what should be considered truly human and intelligent behavior is raised and moved to other domains. We see - not without some concern - how the area of what we consider unreproducible by machines seems to shrink. One may wonder whether we are now witnessing this narrowing process in the aesthetic field. This raises questions such as: could machines reach a point at which we consider them truly creative? How could machines tackle the conceptual turn in contemporary art movements? What role could they have in helping us to understand "good taste" and "bad taste"? Do systems using data analysis tap in to the "unconscious" structure of our culture, or do we witness the emergence of an entirely new form of cultural production?

The original definition of aesthetics as a philosophical discipline was coined in 1750 by German philosopher Alexander Baumgarten and referred to the ancient Greek *aesthesis*, which means sensation or perception. Kant later redefined the term in his *Critics of Judgment* (1790) as the domain of subjective judgments of taste. This meant that aesthetics treated perception as a more complex notion than just sensory experience (investigated today by the psychology of perception), as it also sought to address our affective and cognitive responses to perception. Machines learn to recognize increasingly complex patterns in data that humans are not able to detect. All this raises the following question: to what extent are machine perception and pattern recognition mechanisms relevant for "aesthetic perception," and what are the typically human aspects of aesthetic sensibility that still need to be tackled by artificial systems?

What unique aspects of human aesthetic sensibility still need to be learned by artificial systems?

How AI relate to aesthetics: a simple map

As we saw, computational approaches to aesthetics cover a wide range of applications, from analysis of cultural artifacts to their generation, dealing with questions such as:

a**) Can we develop systems that extract all relevant features of an artifact or an image?** Can we analyze/describe the aesthetic features of aesthetic artifacts from a given cultural tradition?[8]

From a different perspective, we are also interested in questions such as:

b) Can we use AI to understand (and predict) what people like?

We can see a distinction here between questions dealing with *objects* and questions dealing with *subjects*. Concerning the first, we focus on artifact's formal and expressive features (for example, the style of a painting, its motifs, the organization of shapes and strokes, formal similarities to other works), and their semantics and meaning. On the other hand, when we address questions concerning subjects, we grapple with viewers' aesthetic experience and perception, including judgements of artistic value, appreciation, affective and cognitive reactions, etc.

The objective/subjective pair distinguishes between two completely different perspectives found in computational approaches: the first concerns the analysis of objects and aims to extract patterns and stylistic invariants by starting with large databases of aesthetic artifacts and cultural products. The

subjective analysis asks which properties of an artifact correlate with (and predict) people's aesthetic responses, feelings, and interpretations, both individually and collectively.

There is another distinction to be made. Machine learning is used both to *extract* patterns from data and to *generate* patterns after training with said data. Therefore, developments in these technologies not only allow us to *describe* artifacts and predict people's behavior, they can also be implemented to *generate* artifacts and *simulate* people's behavior. Therefore, other kinds of questions should be added: Can we (re)produce what people like and *generate* aesthetically valuable new artifacts? Can we build computational models of people's aesthetic preferences that will allow us to simulate and automate their judgment?

By crossing the two pairs of dimensions - **object vs. subject** and **description vs. generation** - we can identify four different applications of machine learning and AI in aesthetics:

	Pattern recognition (analysis and description)	**Pattern generation** (production and prediction)
Objects	**Studying objects**	**Generating objects**
Subjects	**Studying subjects**	**Generating subjects**

To illustrate the different fields of this map, let's consider the work of Johann Sebastian Bach. His music has been described as highly structured and mathematical, the "chess of music" so to speak, and has been the object of both algorithmic description and generation (like the project "DeepBach" from 2016):

1) **"Studying Objects":** the AI, using a dataset that contains all of Bach's compositions, analyzes melodic patterns, tracking similarities between different scores and extracting the characteristic style of the composer;

2) **"Generating Objects"**: the AI, having been trained with the dataset of Bach's compositions, is used to generate new Bach-sounding variants.

However, an essential aspect of aesthetic analysis would be missed if either of these tasks (analysis of the formal features of a music composition and the production of variants) failed to consider how people react and experience the music. This is where the issue of the subjects' response comes in:

3) **"Studying Subjects"**: preferences are gathered and analyzed in order to determine which musical features are especially preferred or which musical qualities determine a specific aesthetic reaction (a feeling, a mood, etc): think about how online music platforms algorithmically track user preferences. If variance among individual preferences is not too big, it is possible to build a model of aesthetic evaluation in the domain of Bach's compositions. The model generates *predictions* of how a user would evaluate the new Bach's chorales. In turn, listeners hear these new compositions and provide the model with further feedback. If the variance of user reactions is too big, we can use cluster analysis to identify different types of preferences and generate different models that are suitable for each type. This approach would not be dissimilar to companies that "segment" their market's customers into smaller groups based on demographics, interests, needs, behaviors, and/or location. In fact, describing and predicting people's aesthetic behaviors based on previous listening choices constitute the evolution and refining of traditional consumer preference analysis as marketing and sociological research practice. Contemporary approaches, however, use data in a way that affords new analytic capabilities. While traditional market and sociological surveys

typically pool data, use aggregate statistical averages, and form clusters based on theoretical sociological models of human types, algorithmic tracking and analysis of data are capable of generating personal profiles that use individual behaviors as data, such as clicking or liking particular images on a social network or listening to specific music on Spotify or Youtube. Rather than clustering data from many subjects, each profile is unique to one individual.

4) "Generating Subjects": Recommendation systems on online platforms use models that predict what a user would appreciate. However, by modeling a person's aesthetic judgment, it is also possible to *generate* behavior and judgment. Modeling listeners' preferences and aesthetic responses enables us, in principle, to simulate how people would behave and react in front of specific objects. If a composer (or the AI itself) were to create a new variant of Bach-like music, an artificial system trained according to a subject's aesthetic model could formulate evaluations on its own without needing to refer to a human subject.

It is not hard to imagine that "artificial judgment" systems might be increasingly used in the future. These systems would autonomously evaluate cultural objects, scoring a design artifact, fashion item, or image with a higher or lower aesthetic value. An artificial judge could do more than tell us "what we may also like" (as in traditional recommendation systems). It could also tell us "how much people would appreciate" a specific aesthetic artifact that has been submitted to the system, how people would judge it, even predicting what people would tell us about it.

Automated systems for predicting image aesthetic score are a typical example of artificial judgment. These function by using a combination of objective metrics (image quality, sharpness, optimal contrast, colors, etc.) and subjective evaluations. To create such a system, large numbers of people rate lots of images. This data is then used to train a neural network, which can subsequently rate new images automatically.[9] Moreover, we can add that these algorithms could be able to identify aesthetic properties (on the side of ob-

jects) and individual preferences (on the side of subjects) of which people are not even aware, but that are manifested in their appreciative behavior.

Patterns of explanation, or what do we do when we talk about aesthetics

Computational analysis may enable us to extract patterns and formal structures, but it does not provide an understanding of how such patterns affect human perception, emotion, and cognition. Patterns considered in isolation from human meaning are ultimately empty. Art historian Michael Baxandall (in his 1985 book *Patterns of Intention*) has persuasively described the essence of the critical language that we use when talking about any artwork or cultural product. For Baxandall, any discourse that we create is neither a merely factual description of features, nor a subjective report of a person's reactions, but consists in highlighting the *relationship* between the object and human responses (the meaning they give and the aesthetic reaction they manifest). This relationship is further mediated by an understanding of the object's symbolic and cultural meanings. A critic, so to speak, tells the reader what kind of reaction is expected (or would have been expected for people in the past) in front of a specific object. Expressed in the terms of the map from the previous section, this would mean drawing a connection between the description of the object and the description of the corresponding subjective reactions.
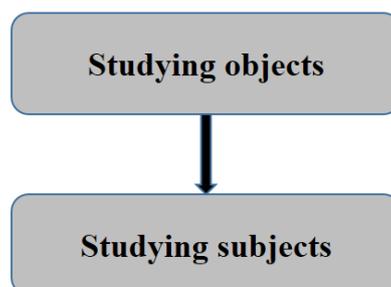
Following Baxandall, cultural and critical explanations in art are not mere descriptions or classifications: they are "primarily a representation of our thoughts about it" (*Patterns of intention*, p. 10). What we describe is a "partially interpretative description": "one does not describe pictures, but our thoughts of having seen pictures" or at least hypotheses on those thoughts. The efficacy of a critic's argumentation lies in his or her ability to compellingly persuade the reader that the artifact elicits the kind of reactions and thoughts that the critic is claiming to make explicit. Moreover, the critic's use of words and concepts, while sharpening the perception of an object, at the same time deepens the meaning of the concept itself: "concepts and object reciprocally sharpen each other" (*Patterns of intention*, p 34). For example, if we describe *The Scream* by Edvard Munch (1893) as inspiring a sense of *dread*, then the very concept of dread as an aesthetic notion will be made richer by using Munch's famous painting as a case in point.

Artifacts in synthetic media (images, songs, texts) are generated by networks that have already been trained on large databases of similar, preexisting artifacts, as in following diagram:

However, if the generated content is expected to have aesthetic value, the generative networks must take into account not only the formal dimension (*how* artifacts are made), but also their corresponding subjective interpretation and reaction, including people's aesthetic preferences. Otherwise, we would be able to generate infinite variations of patterns, but not have a clue on how they relate to our appreciation. If description of patterns without meaning is empty, as we said, generation of patterns without human interpretation is blind.

In AI-media generation today, humans operate generative networks by selecting, adjusting, and tweaking the process to obtain a desired result. This result also depends on humans following their own aesthetic sensibility: for example, a music expert had to evaluate and filter the different generations of Schubert's "Unfinished Symphony". Only algorithmic analysis of subjective responses ("Studying subjects") would allow a progressive automation of this evaluative step.
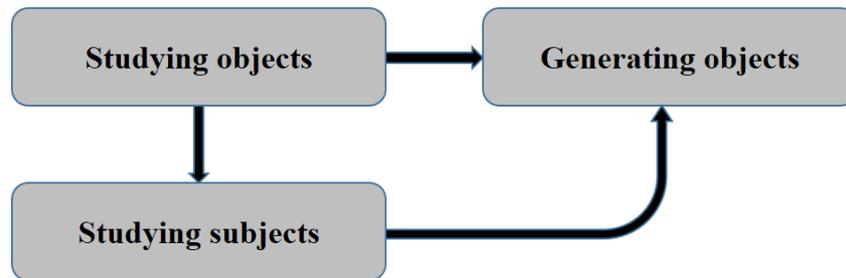
Moreover, "Studying subjects" would involve both individual and collective reactions. The latter involves analyzing historically sedimented responses towards cultural objects. Ideally, an AI capable of creating meaningful art and design would take into account the history of what exists, not only to extrapolate patterns from the artifacts, but also to interpret their collective reception, that is, how people over time have reacted to these artifacts. Thus, **using an AI to generate new cultural artifacts (and assist human creators) will require using an AI for cultural analysis**. This would necessitate bringing artificial aesthetics into contact with the various fields that deal with this issue: philosophical aesthetics, art history, psychology of art, anthropology and sociology of culture, and so on. Granted, new technical developments can generate entirely new kinds of artifacts that need not resemble the cultural production of the past. However, if we want to better grasp how these artifacts could affect people, an understanding of how we typically react and give meaning to aesthetic objects could save us from wandering in the dark. The near future may hold entirely new aesthetic artifacts, but it is unlikely to hold an entirely new human nature.

Computation and psychology

Aesthetic phenomena involve a complex relationship between all human faculties, from low-level perceptual mechanisms to higher-level affective and cognitive processes. It is no coincidence that by the end of the 19th century Gustav Fechner, the father of experimental psychology, had already identified aesthetics as the most critical challenge for his new methods in scientific psychology.[10] In fact, researching how people react and behave during an aesthetic experience (the domain we defined as "Studying people") has long been a tradition in so-called psychological experimental or empirical aesthetics. Fechner investigated, for example, whether people prefer shapes that follow the golden ratio rule. While Fechner's findings seem to confirm the rule, later studies failed to replicate the same results. This line of research continued steadily for almost a century: for instance, Birkhoff's *Aesthetic Measure*[11] tried to capture in a quantifiable formula the optimal aesthetic relationship between a shape's complexity and order: high order with high complexity would correlate, according to him, to a higher aesthetic pleasure. In the 1970s, Daniel Berlyne's new experimental aesthetics[12] introduced motivational factors as a key component in aesthetic pleasure and appreciation: aesthetic value is not only a function of an object's features, but also of the hedonic tone of a subject, namely his or her level of interest and stimulation. His inverted-U relationship between complexity and enjoyment suggests an optimal middle point between too little and too much complexity in a stimulus. This has been empirically investigated as well, albeit with divergent results. At the turn of the new century, researchers felt that it was necessary to move from aseptic psychophysical experiments based on simple abstract patterns to observing how people react in front of real artworks, artifacts, or natural entities.[13] Neuropsychological approaches have recently become popular in this field, extending their focus to issues such as creativity and the mechanisms of reception and interpretation in specific art forms (visual artworks, music, movies, literature).

A researcher in this field typically conducts experiments with small groups of people under carefully controlled conditions, using statistical techniques to analyze the collected data. For example, in many experiments in visual aesthetics, a group is shown a particular set of images (the dataset can be pre-existing or created specially for the experiment), and people are asked to express their preferences in some way, such as rating all images on a numerical scale. Decades of investigation in experimental aesthetics led to many findings. For example, psychologists showed that more prolonged exposure to a stimulus leads to a growing familiarity with the object, inducing a preference for it as well as for prototypes in the object's category. That is, we like what is more typical, and that overall fluency, the ease in processing an experience, correlates with aesthetic preference. Furthermore, research findings showed a preference for symmetry in facial features, a preference for smooth and curved shapes over angular ones; specific preferences for natural landscapes over man-made scenes, and for architectural scenes with naturalistic aesthetics.[14] Numerous studies have tested the classical rules of harmony, balance, and "good composition," such as the "rule of thirds" or the principles described by Gestalt-theory (which were first applied to art by Rudolf Arnheim in his 1954 classic work, *Art and Visual Perception*).

We should note that these experiments often use college students as their test subjects. Their aesthetic judgment could mirror a specific taste, without being representative of the judgments of artists, designers, or critics. Different studies have repeatedly confirmed a significant difference between experts and non-experts in aesthetic evaluation. It should be noted, moreover, that most of the research does not point to conclusive findings, showing instead that aesthetic preference depends on numerous underlying variables, like context and subjective attitudes. One example of a contextual factor would be the verbal description of an artwork: titles change our appreciation of paintings and how we look at them.[15] The order of presentation (which object do we see first? Which next?), spatial disposition (which object is on the left? Which on the right?) and juxtaposition (do we compare similar or

---

Big data does not require us to assume a universal human aesthetic subject

very different objects?) also affects how people judge objects.[16] The environment also influences how we evaluate and appreciate art.

For instance, our reception of an artwork may differ depending on whether we look at it in a typical "white cube" space or in a more informal context. Variations exist depending on the observer's characteristics: factors such as one's emotional state and level of arousal, expertise, personality traits and culture all contribute to the aesthetic experience and judgment. Instead of looking for generic universal rules – like the golden ratio, "unity in multiplicity," and Berlyne's inverted-U model – experimental research investigates very subtle mechanisms while considering contextual, personal, and culturally specific factors. In summary, the field has generated and tested many interesting theories to account for human aesthetic experiences, demonstrating at the same time that none of them seem to hold universally.[17]

There are two crucial differences between today's computational methods and traditional experimental aesthetics. First, experimental aesthetics mostly focuses on subjects, while artificial aesthetics focuses on objects. Furthermore, experimental aesthetics uses specially selected and highly controlled stimuli, while artificial aesthetics uses "big data" from real life human behavior, which is often collected through digital platforms.

While experimental aesthetics usually produces stimuli in controlled settings and looks at people's responses, computational methods make use of large, available datasets of expressed preferences, like Photo.net or Dpchallenge.-com (used for computation studies in the late 2000s), allowing researchers to explore how people give their "likes" on social platforms. In other cases, they capture and measure people's actual consumer behavior on online platforms,

like streaming services for music and film, with the aim of inferring features from the most popular artifacts.

In experimental aesthetics, a subject-focused approach emphasizes the analysis of so-called "dependent variables". These include the controlled responses of subjects, measured through judgments on well-calibrated scales, as well as physiological reactions (heart rate, skin conductance, pupil dilation etc.) and brain activity, measured with EEG or fMRI, which theoretically obviate the problems associated with verbal evaluation. Computational analysis of aesthetic behavior, on the other side, is an object-focused approach and is particularly strong at describing "independent variables", i.e., the aesthetic contents that are consumed and judged by people every day. This strength stems from its capacity to gather and analyze large numbers of features from images, music and other cultural artifacts. As previously mentioned, the key advantage of computational approaches to aesthetics is the fact that they are not bound to seek aesthetic universals or to take the common responses of (relatively small) groups of subjects to be representative of general attitudes. Instead, algorithms can track individual preferences and behavior without needing to model aesthetic responses based on aggregated averages. **Big data does not require us to assume a universal human aesthetic subject.**

Despite these advantages, an artificial aesthetics that focuses on aesthetic preferences still has to deal with the methodological challenges that characterize all experimental approaches. We shall briefly mention two of them, concerning 1) the difficulty of isolating the features linked to our aesthetic evaluation, and 2) the difficulty of determining what kind of response we are trying to describe.

Concerning the first point, features of aesthetic objects are hard to isolate. For example, to study how variations in the shape of a design item influence aesthetic appreciation, an experiment should use a controlled setting that analyzes the effect of minimal variations in the shape and avoids confounding multiple variations at once (e.g., changing shape and color, or shape and texture etc.). However, aesthetic variables can also interact with each other.

Consequentially, this set-up would not allow us to draw a one-to-one correspondence between the feature and the aesthetic responses to the feature on this particular object. It is certainly possible to determine general trends in people's preferences: e.g., we could observe that a certain musical style is more popular than another one with a particular demographic in a given country. However, it is not always easy to reach greater granularity and comprehend the precise role of each factor in the final aesthetic effect: what exactly makes the one musical style more appealing than the other? In order to achieve this level of understanding, we would need a large number of similar aesthetic artifacts that present only small variations from each other.

In some cases, digital platforms allow us to study a vast number of different but not too heterogeneous stimuli which are available on the web. For example, in a study from 2014[18], the authors used hundreds of features from micro-videos (up to six seconds' duration) on Vine, a former media sharing platform, to predict whether people would judge them as "creative" or "non-creative". The study used a crowdsourcing platform to have 284 people judge 3800 videos. Each video received evaluations from multiple people, the average agreement of which was calculated to be 84%. The features covered scene content, filmmaking techniques, photographic techniques, composition, visual affect, audio affect, and novelty. All these features were defined mathematically and calculated automatically from the videos through an analysis of their frames and soundtrack. The authors report the classification accuracy for each group of features, concluding: "The best results are achieved when we combine novelty features with aesthetic value features, showing the usefulness of this twofold definition of creativity." Used separately, composition and photographic techniques outperform scene content (classification accuracy is 77% vs 73%), while novelty video features outperform novelty audio features (74% vs 63%). To get these kind of results, it is necessary to have a sufficiently wide data set whose features are manageable (like a short six-second film), which is not always the case with human cultural production.

23

What are we actually measuring when we ask a subject about her aesthetic experience?

Concerning the second point, human aesthetic responses (i.e., dependent variables in a psychological experiment) also pose their own challenges. What are we actually measuring when we ask a subject about her aesthetic experience? Our relationships with aesthetic objects have many layers and dimensions. They can range from sub-personal physiological reactions to complex critical formulations, from a "like" given to an image in a social network, to actual consumption behavior, up to sophisticated critical judgment. We get a different answer depending on whether we ask someone if she "likes" a movie or if she considers it a masterpiece, or if we simply observe her physiological reactions while watching that movie. Moreover, we should distinguish between value judgement and mere subjective preference/desire: in general, we can say that value judgments are more stable than momentary preferences or desire for a certain object. I can consider song X to be a masterpiece (and superior to song Y), but lack the desire to listen to X at present, instead experiencing a greater desire to listen to Y, maybe because of my emotional state or because I listened to X too many times. This means that my consumption behavior can reveal preferences that do not necessarily express my general idea of aesthetic value: I may be an avid consumer of action movies and yet consider arthouse films aesthetically superior, even though I watch them more rarely. Artificial systems that gather data about human aesthetic consumption should take these issues into consideration if we want to avoid overly simplistic models of human aesthetic experience and judgment, both of which are used in artificial evaluative and generative algorithms.

[1] Datta et al. propose 56 different rules and features. Datta, Ritendra, et al. "Studying aesthetics in photographic images using a computational approach." *European Conference on Computer Vision*. Springer, 2006, pp. 288-301.

[2] Kao, Yueying, et al. "Deep aesthetic quality assessment with semantic information." *IEEE Transactions on Image Processing*, vol. 26, no. 3, 2017, 1482-1495. *IEEE Xplore*, https://doi.org/10.1109/TIP.2017.2651399.

[3] Hadjeres, Gaëtan, et al. "DeepBach: a Steerable Model for Bach Chorales Generation." *Proceedings of Machine Learning Research*, vol. 70, 2017. *arXiv*, https://arxiv.org/abs/1612.01010; Emerging Technology from the arXiv. "Deep-Learning Machine Listens to Bach, Then Writes Its Own Music in the Same Style." *MIT Technology Review*, 14 December 2016, technologyreview.com/2016/12/14/155416.

[4] "Can You Tell the Difference between AI and Human Composers?" *YouTube*, uploaded by TwoSetViolin, 23 September 2020, youtu.be/PmL31mVx0XA; "Bach vs AI: spot the difference." *YouTube*, uploaded by Orchestra of the Age of Enlightenment, 22 November 2019, youtu.be/lv9W7qrYhbk.

[5] Davis, Elizabeth. "Schubert's 'Unfinished' Symphony completed by artificial intelligence." *Classic FM*, 6 February 2019, classicfm.com/composers/schubert/unfinished-symphony-completed-by-ai.

[6] Xue, Alice. "End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks." *Proceedings of IEEE WACV*, 2021, pp. 3863-3871. *arXiv*, https://arxiv.org/abs/2011.05552.

[7] By using well-defined and ill-defined problems, we are referring to the crucial distinction made by Herbert Simon in 1973 regarding artificial intelligence. See Simon, Herbert. "The structure of ill structured problems." *Artificial Intelligence*, vol. 4, 1973, pp. 181-201. *ScienceDirect*, https://doi.org/10.1016/0004-3702(73)90011-8.

8 The practical applications of such analysis include designing interactive online interfaces for museum collections. For example, when a person chooses a particular artwork, the system shows other artworks in the collection that are most similar. Further applications can be found in digital art history: changes in any extracted feature or a combination of features can be plotted over time to analyze the evolution of a single artist or entire historical periods.

9 See Djudjic, Dunja. "The Rise Of The Machines: Google's AI Will Decide If Your Photos Are Aesthetically Pleasing." *DIY Photography*, 26 December 2017, diyphotography.net/rise-machines-googles-ai-will-decide-photos-aesthetically-pleasing; and Mikhailiuk, Aliaksei. "Deep Image Quality Assessment." *towards data science*, 15 March 2021, towardsdatascience.com/deep-image-quality-assessment-30ad71641-fac.

10 Fechner, Gustav. *Vorschule der Aesthetik*. Breitkopf und Härtel, 1876.

11 Birkhoff, George D. *Aesthetic Measure*. Harvard University Press, 1933.

12 Berlyne, Daniel E. *Studies in the New Experimental Aesthetics*. Wiley, 1974.

13 Leder, Helmut, et al. "A Model of Aesthetic Appreciation and Aesthetic Judgements." *British Journal of Psychology*, vol. 95, 2004, pp. 489-508.

14 For an overview, see Locher, Paul L. "Contemporary Experimental Aesthetics: Procedures and Findings." *Handbook of the Economics of Art and Culture*, edited by Victor A. Ginsburgh and David Throsby, vol. 2, North Holland, 2013; *The Cambridge Handbook of the Psychology of Aesthetics and the Arts*, edited by Pablo P. L. Tinio and Jeffrey K. Smith, Cambridge University Press, 2014.

15 Leder, Helmut, et al. "Entitling Art: Influence of Title Information on Understanding and Appreciation of Paintings." *Acta Psychologica*, vol. 121, 2006, pp. 176-198.

16 Khaw, Mel W., and David Freeberg. "Continuous Aesthetic Judgment of Image Sequences." *Acta Psychologica*, vol. 188, 2018, pp. 213-219.

[17] Even the assumption that beholders would universally associate particular forms with specific qualities or "aesthetic effects" has been put into question. See Specker, Eva, et al. "Warm, Lively, Rough? Assessing Agreement on Aesthetic Effects of Artworks." *PLOS One*, 13 May 2020, https://doi.org/10.1371/journal.pone.0232083.

[18] Redi, Miriam, et al. "6 Seconds of Sound and Vision: Creativity in Micro-videos." *IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 4272-4279, https://doi.org/10.1109/CVPR.2014.544.