When scientists fail to look at Mona Lisa they see numbers but not the art

Mihai Nadin
Institute for Research in Anticipatory Systems
University of Texas at Dallas
www.nadin.ws
Corresponding: nadin@utdallas.edu

Abstract: Computers as media for representation are omnipresent. The advantages of digital representations ought to be understood beyond convenience. Such representations entail all the implicit assumptions of computation and all its advantages. They falsify—as does every medium—the represented, offering in high resolution the illusion of reality. We use an example to empirically demonstrate that the expectation of experimental reproducibility is eliminated in the computational representation. Given the fact that as syntactic processes characteristic of computation do not support the semantic or the pragmatic dimension of representation, no inferences to the represented can be made.

Keywords: reproducibility; representation; data; information; meaning; anticipation

In a previous study (IJARITAC, 2:1, 2011), we focused on the relation between computation, data, information, and meaning. A recent article (Horstman and Loth, 2019) inspires a more focused return to the subject.

Reality, to which we belong as much as whatever we try to understand, is taken in by the individual through the senses. Machines can also take reality in, via artificial sensors, which emulate our senses. Digital photography, video, multimedia representations, visualizations, and simulations are examples we encounter on a regular basis.

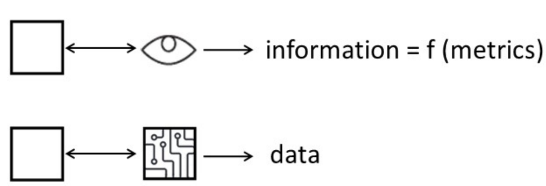


Figure 1. The perception of reality vs. the measuring of reality. The living harvests information; machines collect data.

Imagine the following: leaves are falling. The data view would focus on how many leaves, frequency, timeline, temperature, humidity, etc. The outcome would be data, and the more we wish to have, the more we have to measure. Interpretation of data returns information about the process. Researchers in botany would reference the data for their significance: Do leaves fall once the temperatures are getting low? Or do they fall in anticipation of the freeze?

This short paper is not about botany, but in reply to Horstman and Loth (2019), who measured how some undefined set of viewers perceived Leonardo da Vinci's famous painting, the *Mona Lisa*. The authors condensed the result of what they called an experiment in the title: *The Mona Lisa Illusion—Scientists See Her Looking at Them Though She Isn't*. In this article we are interested not in art history, not even in the dubious discipline of psychology (where the authors seem comfortably ensconced), but rather in the cognitive discipline of representation, as a prerequisite for understanding reality. The following are the arguments for questioning both the so-called experiment conducted by the authors and the conclusion they reached:

1. The authors do not provide any proof that the perception of the original (in this case a famous painting) is the same as, or in some ways equivalent to, the perception of photographs or scans of the original. In other words, they do not distinguish between the represented and the representation.
2. The authors do not provide proof that size and/or zoom of the representation (the digital facsimile of the original work) have any impact on the viewers' perception.
3. The authors do not even attempt to justify the implicit assumption that characteristics of the subjects (in particular age, culture, gender) affect the outcome. Cognitive bias is of no concern to them.
4. In respect to the visual perception involved, the authors do not distinguish between static and dynamic gaze.
5. Relevant references, in particular regarding the role of meaning, are missing. (But the lack of reference discipline—a subject that should worry the scientific community—is not their only shortcoming!)

To keep the core idea of this text short: This is yet another example of an experiment that cannot be replicated due to the nature of the representation chosen by the researchers. In what follows, I shall address each of the aspects mentioned above. A contrasting empirical evaluation will be provided with the aim of showing that the reductionist approach—reduction of dynamic to static visual perception is only one part of the larger subject—results in experiments that are not reproducible.

The issue of representation

When the real (the *La Joconde* portrait in the Louvre) is replaced by a representation (photo, scanned image, a painted copy, a print, etc.), two aspects ought to be kept in mind: 1) no representation is complete; 2) in the process of transposing the original onto a substitute medium, (in this case an image on a computer monitor), noise is introduced, and the image itself is changed. While the authors provide details regarding the scanned image (7,479px by 11,146px) and the monitor used (a 35 x 26 cm computer screen), they fail to assess the way in which the scanning and display technology affected the characteristics of the image as it was transposed from an oil painting to a display on a computer. They use a re-presentation of the original, never questioning the impact of the technology on the representation. It is evident that the scanned image displayed on a computer monitor (at 66 cm distance from the viewer) affords a perception different from that of the original, or at least from what generations of researchers (art historians, psychologists, vision scientists, etc.) have examined before formulating the Mona Lisa gaze effect.

To represent is to present again. Figure 2 consists of the reproduction of the original—the framed painting hanging on a wall in the Louvre—and a small subset of reproductions: books about the painting, music inspired by Mona Lisa, the original name (*La Gioconda*), the name under which it is displayed (*La Joconde*), the experiment itself (which is in its own way a representation), the various zoomed edited variations mentioned in the article, etc. Obviously, the music inspired by the famous painting will not smile at us. This is an open-ended selection of representations, never to be complete, as no representation can ever fully re-produce whatever it represents. Each representation is pragmatically defined: the purpose pursued informs the choice. If a number, such as the weight, or the price of the work, would have been used, it would be telling about how strong should be the box in which the work can be transported, or how much money is to be expected in selling the painting.

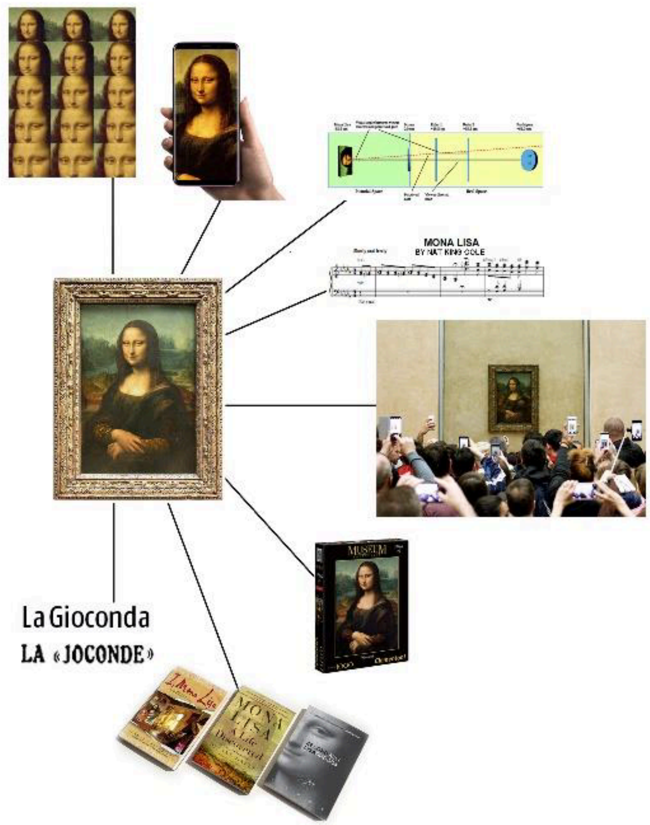


Figure 2. Possible representations of Mona Lisa—open-ended list of possibilities—including the experimental setting used in order to measure whether La Gioconda (the name is also a representation) is gazing at the viewer. The setting for the so-called experiment is easy to identify (above the musical score).

The sample

The gaze-metric—cleverly referenced by a simple carpenter's rule—of an image on a computer monitor might be independent of the cognitive profile of the 24 participants, but it does not afford any insight into dynamic perception. The perceptual characteristics of the 24 subjects, never mind their cultural profiles, are left out. The authors have not proven that this sample is relevant or that the characteristics of the subjects are relevant. As I learned from them (via the journal *iPerception*), they expect the reader to carry out this research for them: Is the sample relevant or not?.

The authors wanted to test the gaze in order to prove a point. Direction of gaze and of the artificially defined line of gaze are not the same. It is obvious from their test setting that they worked with a reductionist model of gaze perception (Garner & Hecht, 2007). Elements specific to the problem at hand are left out, in particular the head position of the viewers and their vision characteristics (perfect vision is taken for granted!). It is because of this failure to ground the problem that is the subject of their experiment that the authors do not understand that viewing an image on a computer screen from a distance of 66 cm is different from the perception of a painting hanging in a museum. Viewers get closer in order to discern details, or seek distance (of an individually based action) in order to realize the holistic nature of the image they examine. The background against which a painting is displayed, or the lighting, even the sounds in the museum, affect our perception. No one has ever claimed that *reproductions* of Mona Lisa look at us or smile. The fact that the experiment left out "non-perceptual information including beliefs"—which means whether the viewers even knew what they were looking at, never mind if they studied art—is a reduction within the mechanistic view they adopted.

Among the missing representations are quite a number of attempts to explain why the Mona Lisa's eyes follow you (e.g., David Munger's entry from March 2005 in the *Cognitive Daily*). The references to those who have dealt with the subject are incomplete (Reibe, DiPaola, Enns, 2009; Quiroga & Pedreira, 2011). Missing are the attempts of art historians to explain gaze—in particular to define the characteristics of linear perspective—and the aesthetic considerations. Within information aesthetics (as limiting as it was), attempts were made at quantifying a variety of characteristics of aesthetic artifacts (Nadin, 2009; Nake, 2012).

Distinguishing between data, information, meaning

Sensory evidence (what we see, hear, smell, etc.) results in data. Sensorial perception (using sensors to acquire from reality the results of measuring) results in data. If a person or a machine interprets the data, information is extracted. Finally, if we interpret reality (in context), we access meaning—never reducible to quantities (Nadin, 2018a).

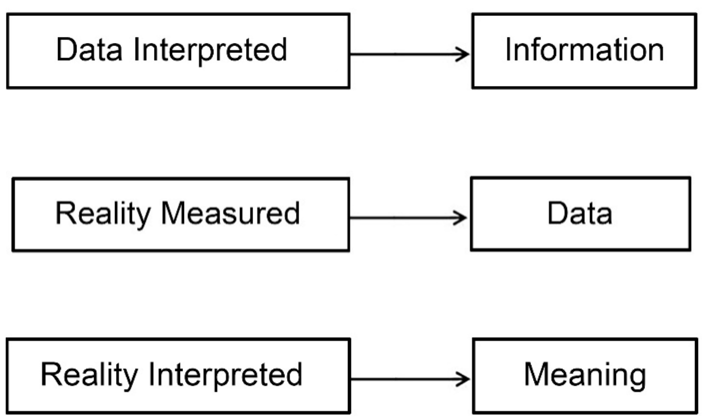
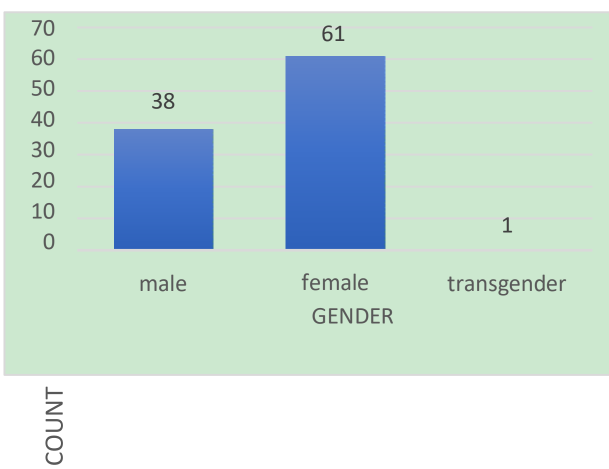


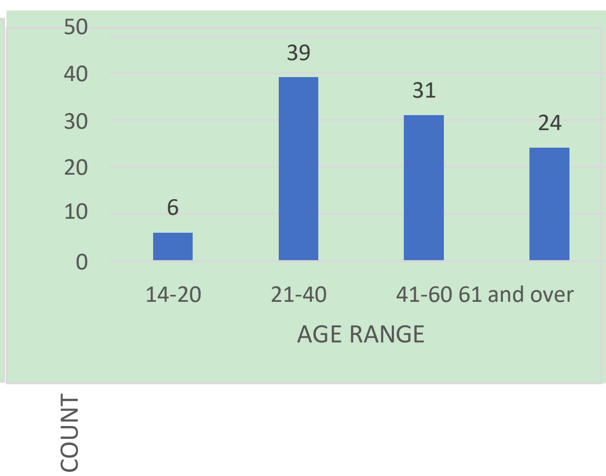
Figure 3. Data represent quantitative descriptions of reality. Data associated with reference (what the data are about) provide information about the observed. Reality interpreted (in respect to a goal, such as a task to be performed) supports the constitution of meaning.

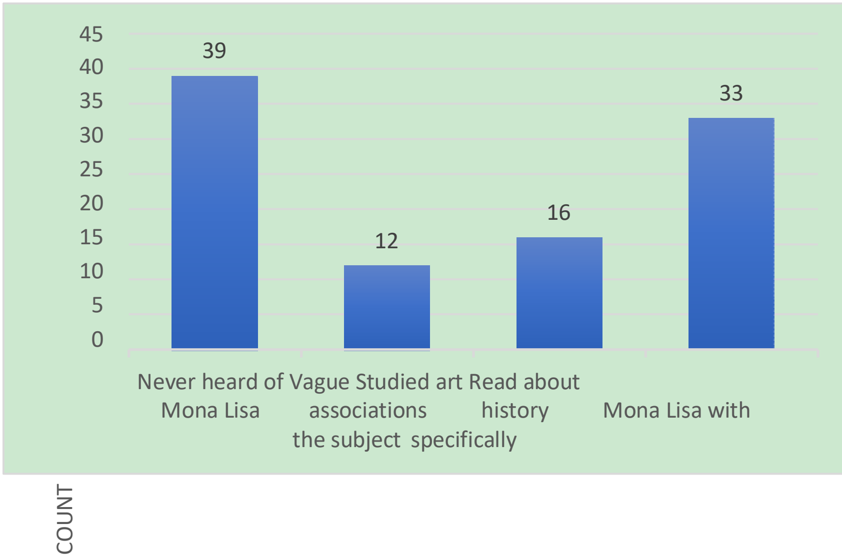
Mona Lisa on the monitor is data represented by a configuration of pixels. It could just as well be a matrix of numbers—depending on the system used, e.g., digital, hexadecimal, etc. In association with our reference (such as color tables), the data become information. For the computer, as a medium, *smile*, as defined in culture, does not exist; or at best it is yet another collection of data. *Smile* is made into actionable data by ontology engineers (they produce dictionaries for machines). Therefore, to pose a question of meaning (what does the Mona Lisa painting mean?) in reference not to the original hanging in the museum but in reference to a set of data representing the art work, is scientifically as dubious as wondering whether the earth is not flat since we don't fall off it as we move over the globe's surface. Ontology engineers make the "understanding" by the machine called Siri (a program for speech recognition) of a sentence such as *Mona Lisa smiles at us* possible. But neither Siri nor any machine knows what a smile is (never mind why we feel happy when somebody smiles at us). Indeed, Siri is "trained" to understand that *smile* conjures "toothbrush" or "You're on candid camera," but not gravitational waves or GAN (generative adversarial network). These are only examples that illustrate why the pseudo-semantics injected by ontology engineers into the machine learning on which Siri is based does not make the machine understand what is processed.

Empirical findings using an interactive setting

On David Munger's closed blogsite (2005), an academic reference of anecdotal significance, Madison G. (2009; cf. Munger, 2005) came up with an idea that I decided to pursue: Mona Lisa as a high-resolution screen saver. One hundred subjects, from ages 14 to 86, of course with variable vision characteristics (from 20/20 vision to corrected vision) looked at the Mona Lisa scaled-down image.







Tables 1, 2, 3: Profile of subjects in the empirical study

The distance was between inches and an arm's length. The question was: What strikes you while examining this image on the iPhone? Without any exception, the subjects noticed a) her smile; b) that her eyes followed them. Of course, this is empirical observation, not an experiment subject to replication. And, of course, it only proves that the image reproduced on the small monitor looks at whoever looks at her. Exactly what Leonardo tried to achieve.

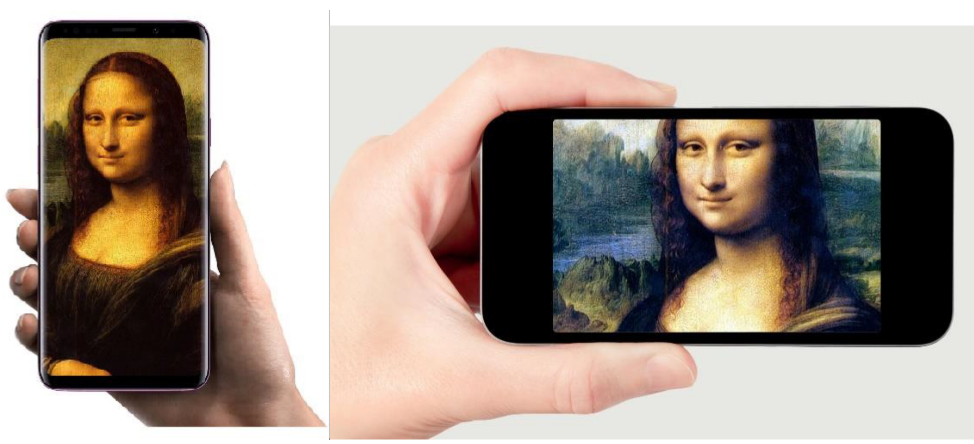


Figure 3. Mona Lisa on an iPhone. Her smile and her gaze noticed by all subjects. She "looks" at whoever looks at her. This is what Leonardo's aesthetics is about.

The nature of the experiment

In this short discussion of a confusing experiment (to which I contrasted empirical findings), I left out more systematic notes on what an experiment is and what the limitations of experiments are (for details, see Nadin, 2018b). Perception is a subject by necessity associated with the living. If artificial perception (as in vision systems, for example) is the object, experiments that replicate can be carried out, provided that the phase space describing the machine supposed to perform the function is well defined. You can prove (like in mathematics or physics) some hypothesis. For living processes, the phase space changes as the perception is influenced by the response (for example, motoric expression—moving closer to the image, squinting, etc.). Since the subject under discussion is visual perception within the knowledge domain of aesthetics, let us provide some details.

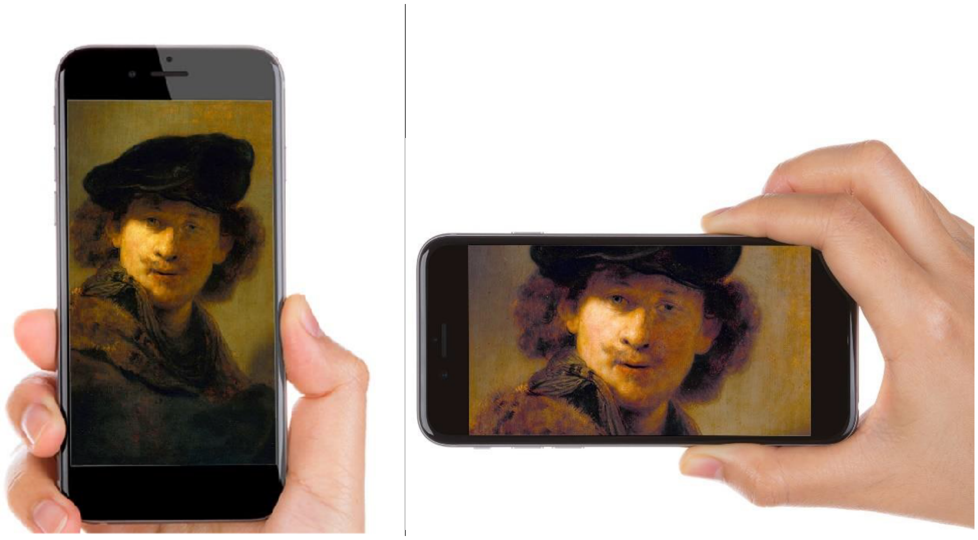


Figure 4. Viewer gaze is informed by various levels of detail. It is the outcome of an aesthetics based on a sophisticated understanding of perspective as dynamic and not static.

The phase space describing visual perception can be reduced, as the authors did in their experiment—the gaze is reduced from its 3D space (the gazing cone) to a 2D situation, but only at the price of losing meaning, which is the definitory aesthetic characteristic. Gaze and gaze direction are two different things. Figure 3 exemplifies the understanding of perspective as a dynamic expression (Rembrandt's *Self-Portrait in a Cap and Fur-trimmed Cloak*). Mona Lisa's portrait is an expression of the same. The experiment considered is marginally appropriate for describing the syntactic level—data from the reading of the carpenter's rule that references the deviation from the line of gaze—but entirely misses the semantic aspect of dynamic visual perception. Not to mention the pragmatics: why *should* Mona Lisa's gaze follow us? What counts is the meaning. The phase space of the variables involved in the gaze perception changes continuously in perception in the museum. A variety of interactions among viewers (including dialogs or listening to a guide) contributes to this dynamic.

The reason to write about the pseudo-science of data reductionism

The empirical case does not demonstrate more than does the so-called experiment that Mona Lisa is gazing at us. But it recovers the most important aspect: the interactive nature of aesthetic artifacts. It shows that the reductionist approach—reduce everything to the data of measurement of physical characteristics—is scientifically questionable. The aesthetic domain, scientifically defined by Baumgarten (1750) is that of meaning. Aesthetic artifacts are perceived through interactive experiences. Meaning is complementary to quantitative assessments of aesthetic expression (and of life in general).

The reason to write about this is twofold: to bring to the attention of the scientific community the obligation to define experiments that preserve the coherence of the perspective; and to avoid placing false conclusions in the public domain. After the publication of the so-called *Mona Lisa Illusion,* the media rushed to tell the public that it was wrong in "seeing" that she was gazing at us, instead of saying that a scan of the painting seen on a computer monitor from 66 cm distance does not confirm the expectation. In the spirit of Popper's legacy (1983), we falsified the experiment in a simple empirical setting. When science disseminates questionable results (think about the major subjects of the day, such as evolution, climate change, sustainability, etc.), and, even worse, uninformed conclusions (in disregard of aesthetics,) it does a disservice to society and to its own credibility. Science depends upon a solid understanding of what representation is and how we ought to learn how to interpret our own data.

Acknowledgments

Abraham Moles and Max Bense—advocating information aesthetics—as well as their students (Nake, Franke, Ness, Maser) have served as a sounding board for a long time. Elvira Nadin guided us in the understanding of the subtleties of the meaning of images. She also pointed to aspects of anticipatory expression. Without Dr. Naz's cooperation, the preparation of this short text would have been close to impossible.

References

Gamer, M. & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 705–715. doi:http://dx.doi.org/10.1037/0096-1523.33.3.705

Horstman, G., & Loth, S. (2019). The Mona Lisa Illusion—Scientists See Her Looking at Them Though She Isn't. *i-Perception 10*(1), 1– 5. doi: 10.1177/2041669518821702

Munger, D. (2005, 2006) Why the Mona Lisa's eyes follow you around https://scienceblogs.com/cognitivedaily/2006/07/31/why-the-mona-lisas-eyes-follow-1

Nadin, M. (2009). Anticipation and the Artificial. Aesthetics, ethics, and synthetic life. *Ethics and Aesthetics of Technologies – AI* & *Society (Computer Science), 25,* 103-118. doi: 10.1007/s00146-009-0243-0

Nadin, M. (2018a) Meaning in the Age of Big Data. In: M. Danesi (Ed.) *Empirical Research on Semiotics and Visual Rhetoric*, pp. 86-127. Hershey PA: IGI Global Publishers, 2018. DOI: 10.4018/978-1-5225-5622-0.ch006

Nadin, M. (2018b) Rethinking the Experiment: necessary (R)evolution. *AI & Society 33*(4), 467– 485. doi: 10.1007/s00146-017-0705-8.

Nake, F. (2012. Information aesthetics: an heroic experiment. *Journal of Mathematics and the Arts 6*(2-3) 65-75. doi.org/10.1080/17513472.2012.679458

Popper, Karl (1983) [reprinted 2000]. Realism and the Aim of Science: From the Postscript to The Logic of Scientific Discovery (1st ed.). London and New York: Routledge. *ISBN 0-415-08400-8*.

Quiroga, R.Q. & Pedreira, C. (2011). How do we see art: an eye-tracker study. *Frontiers in Human Neuroscience 5*, 98. doi.org/10.3389/fnhum.2011.00098