Toward Quantifying Ambiguities in Artistic Images

XI WANG, TU Berlin and MIT, Germany
ZOYA BYLINSKII and AARON HERTZMANN, Adobe Research, USA
ROBERT PEPPERELL, Fovolab/Cardiff Metropolitan University, UK

It has long been hypothesized that perceptual ambiguities play an important role in aesthetic experience: A work with some ambiguity engages a viewer more than one that does not. However, current frameworks for testing this theory are limited by the availability of stimuli and data collection methods. This article presents an approach to measuring the perceptual ambiguity of a collection of images. Crowdworkers are asked to describe image content, after different viewing durations. Experiments are performed using images created with Generative Adversarial Networks, using the Artbreeder website. We show that text processing of viewer responses can provide a fine-grained way to measure and describe image ambiguities.

CCS Concepts: • **Applied computing** → **Arts and humanities**;

Additional Key Words and Phrases: Datasets, generative adversarial networks (GAN), image descriptions, text tagging, aesthetics

ACM Reference format:
Xi Wang, Zoya Bylinskii, Aaron Hertzmann, and Robert Pepperell. 2020. Toward Quantifying Ambiguities in Artistic Images. *ACM Trans. Appl. Percept.* 17, 4, Article 13 (November 2020), 10 pages.
https://doi.org/10.1145/3418054

1 INTRODUCTION

"When looking at a picture, one should say that the more associations it can open up the better."
—Pablo Picasso [7]

When confronted with a new image, the human visual system automatically tries to make sense of it [26, 29, 33, 34]. Some images are easy to interpret, but others require more effort, because they are ambiguous, multivalent, or indeterminate. Art theorists have argued that visual art often exploits ambiguity to engage viewers by simultaneously suggesting and concealing the meaning of a work, or by evoking multiple diverse meanings [12], as Picasso suggests in the quote above. Time plays an important role in these theories: some images are confusing at first, but then lead to an "Aha" moment as the subject is recognized [21, 25], whereas images that appear

Authors' addresses: X. Wang, TU Berlin and MIT, Marchstrasse 23, Berlin, 10587, Germany; email: xi.wang@tu-berlin.de; Z. Bylinskii, Adobe Research, 1 Broadway, Cambridge, MA, 02142, USA; email: bylinski@adobe.com; A. Hertzmann, Adobe Research, 601 Townsend Street, San Francisco, CA, 94103, USA; email: hertzman@dgp.toronto.edu; R. Pepperell, Fovolab/Cardiff Metropolitan University, 200 Western Avenue, Cardiff, CF5 2YB, UK; email: rpepperell@cardiffmet.ac.uk.

1544-3558/2020/11-ART13 $15.00
https://doi.org/10.1145/3418054

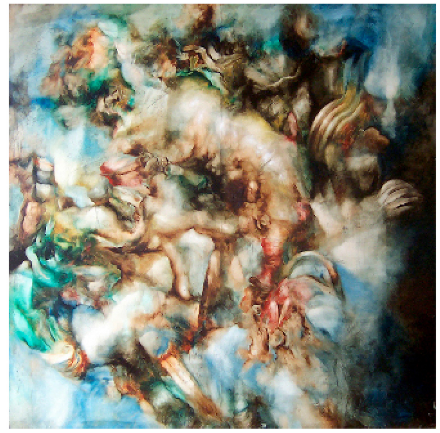


Fig. 1. Two examples of indeterminate paintings by contemporary artists. **Left**: Annunciation After Titian by Gerhard Richter, Oil on Canvas, 1973 (CR 344-1). © Gerhard Richter 2020 (0099). **Right**: Succulus by Robert Pepperell, Oil on Panel, 2005.

simple at first but become more perplexing over time are sometimes called “indeterminate” [27]. Indeterminacy is a major theme in modern visual art [12, 28], e.g., Gerhard Richter is an example of a major contemporary artist who deliberately creates and values indeterminate images [31] (Figure 1).

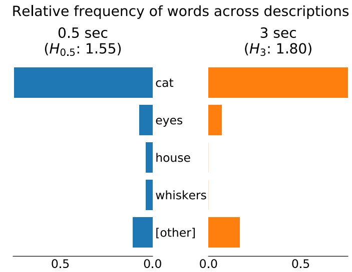
Can we analyze these image properties quantitatively? Several recent authors have attempted to categorize them [22, 23, 27, 28], and to put them in the context of neuroscience theories [15, 35]. However, experimentation remains difficult. Previous studies have tended to use “handmade” artworks [9, 18, 19, 22, 24, 36, 37]. For example, most methods compare artworks made by one artist to works by another artist, or to works by the researchers. Such images have several limitations when used as experimental stimuli. For example, viewers’ judgments may be influenced by historical, stylistic, or contextual factors not of direct relevance to the study. Highly simplified graphics, such as Mooney faces, have also been used as stimuli [21], but generalizing from these results is also challenging. Moreover, previous studies typically ask high-level subjective questions of in-person participants, including whether or not an image is ambiguous or contains an object. As a result, the conclusions from these methods are promising but necessarily limited.

This article proposes an approach to measuring perceptual ambiguity in artistic images. Study participants are shown an image for a fixed duration, and asked to describe the contents of the image. We hypothesized that the quantity and diversity of the descriptions would provide a measure of perceived ambiguity. We also hypothesized that the distribution of responses would vary for different viewing durations, reflecting how perception of an image evolves over time.

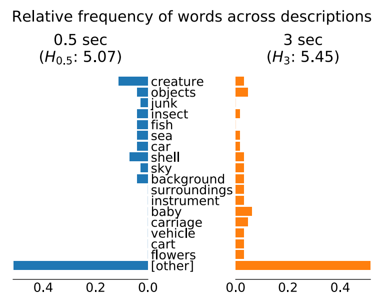
For stimuli, we use images from Generative Adversarial Networks (GAN) [14]. Specifically, we gather popular images from the website Artbreeder.com, since these images span a range of ambiguity and indeterminacy [16]. These stimuli avoid some of the limitations of previous studies: they are presented in their “native” format as digital images, they minimize art historical or contextual confounds, they are visually rich, relatively free of stylistic bias, and can be generated in large numbers.

Understanding ambiguous and indeterminate images is important not only for art theory and history but for our understanding of human image perception more generally. We show that histograms and entropy of viewer response can capture and summarize image ambiguities. Our results suggest how these kinds of studies could help describe, categorize, and measure the space of image ambiguity.

Dur.	Description	Tokens
0.5s	cat	cat
0.5s	cat	cat
0.5s	a white cat	cat
0.5s	grumpy cat	cat
	⋮	
3s	white fluffy cat	cat
3s	a white cat	cat
3s	white cat	cat
3s	cat	cat



Dur.	Description	Tokens
0.5s	animal	creature
0.5s	shell on legs	legs
0.5s	an insect in the sky	insect, sky
0.5s	a fish or sea creature	fish, sea, creature
	⋮	
3s	A calm shell	shell
3s	A robot looking swan	robot, swan
3s	a giant snail	snail
3s	a sea urchin or plant	sea urchin, plant



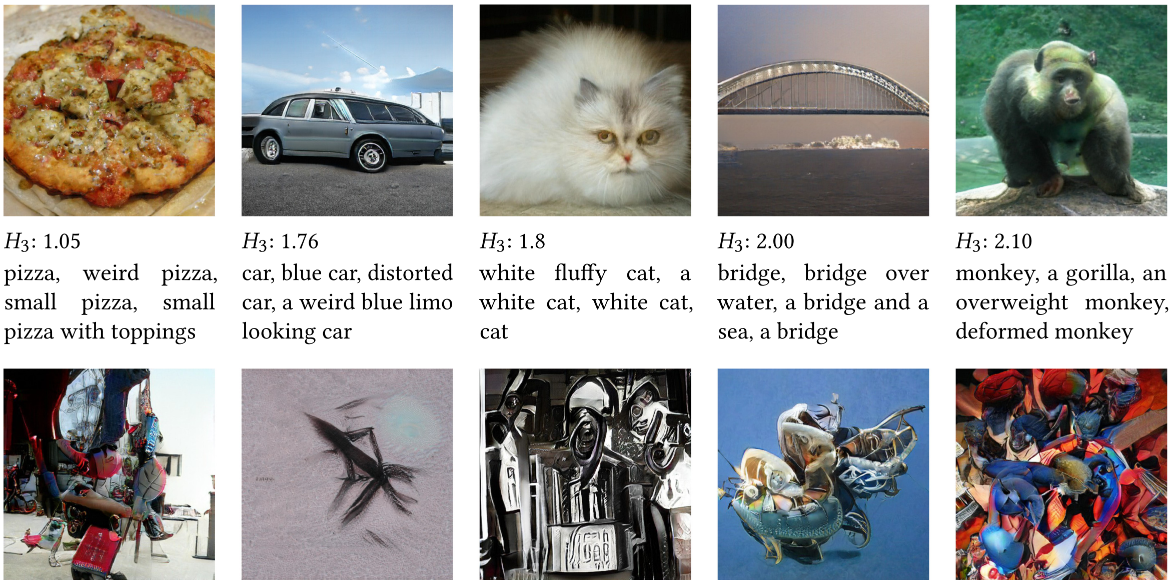
(a) Stimulus image (b) Sampling of image descriptions (c) Description histograms

Fig. 2. Our perceptual study pipeline summarizes the distribution of how participants describe an image after a fixed viewing duration. *Top row:* A sample stimulus image is shown in the upper-left. In the middle column, sample descriptions are shown for each image, along with the text tokens extracted by our processing. Note that the descriptions are quite homogenous for the top row: nearly all participants describe the first image as a cat. This is reflected in the histogram of tokens on the upper right, for both time durations. The entropies are correspondingly low: $H_{0.5} = 1.55$, $H_3 = 1.80$. The "[other]" histogram bin counts tokens that appeared only once each. *Bottom row:* A more indeterminate image yields much more variability in descriptions, and the variability increases considerably over time: $H_{0.5} = 5.07$, $H_3 = 5.45$.

2 PERCEPTUAL STUDY AND DATA PROCESSING

Image stimuli: To form the image dataset, we manually identified a set of 150 images from Artbreeder.com that appeared to exhibit variations in image ambiguity. All images were taken from Artbreeder's "General" class, which provides images from the BigGAN model [5]. The first 120 images were manually selected from among the most popular images on Artbreeder, and loosely categorized into 4 categories of 30 images each: "Recognizable" (e.g., Figure 2 (top)), "Dichotomous" (depicting two or more distinct objects simultaneously, e.g., Figures 7(d) and 7(e)), "Indeterminate" (open to multiple interpretations, e.g., Figure 2 (bottom)), and "Abstract" (clearly containing no objects, e.g., Figures 7(a) and 7(f)). We manually constructed the remaining 30, "AbstractFlat" (e.g., Figures 7(g) and 7(h)) to be highly abstract, using the site's "gene editing" interface, by increasing the BigGAN truncation parameter via the "chaos gene" control and setting all presented embedding coordinates to −1.

Task design: We crowdsourced descriptions using Amazon's Mechanical Turk. A task consisted of a sequence of 30 images, all from a single category of the stimuli set, to avoid confounding effects. Participants viewed each image for either 0.5 or 3 s; the viewing duration was fixed for a single participant, but randomized across participants. After the image disappeared, participants completed an attention vigilance task whereby they reported the last location on the screen where they looked (based on a similar task design [11]). Then, participants entered



H_3: 1.05
pizza, weird pizza, small pizza, small pizza with toppings

H_3: 1.76
car, blue car, distorted car, a weird blue limo looking car

H_3: 1.8
white fluffy cat, a white cat, white cat, cat

H_3: 2.00
bridge, bridge over water, a bridge and a sea, a bridge

H_3: 2.10
monkey, a gorilla, an overweight monkey, deformed monkey

H_3: 5.37
mechanics garage, a mangled person, statue of technology, parade costume

H_3: 5.38
fishing bait hook, clippings from a haircut, chinese writing, black feathers

H_3: 5.43
steel work, metal suit of armor, three tin men hanging out, a radiator

H_3: 5.45
a robot looking swan, musical instrument, a sea urchin or plant, a weird looking seashell

H_3: 5.70
flower buds, stained glass, some sort of bunched up cloth, autumn leaves

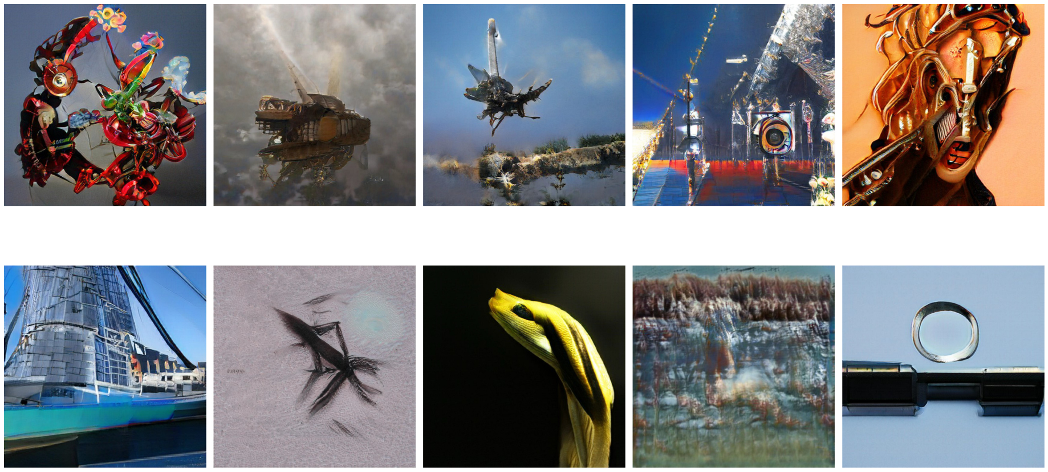
Fig. 3. The images in our dataset with the lowest and highest entropy (H_3) of descriptions. Some sample descriptions are included. Observe that entropy appears to reflect image ambiguity/recognizability, and very indeterminate images have the highest entropy.

a freeform text description. They were instructed to describe the scene and any objects in the image, "even if the image looks abstract at first." We recruited 70 participants for each of the two viewing durations and 5 categories from our stimuli set (launching 700 tasks in total). We filtered participants based on the attention vigilance task. After filtering, we have on average 20.4 descriptions per image and viewing duration. Sample images and descriptions are shown in Figures 2(a), 2(b), and 3.

Viewing duration: Human perception studies have shown that retention of visual details plateaus by 3 s of perception [3, 4], while half a second allows for most visual information to enter conceptual working memory without interference [10, 30]. We ran a pilot study with viewing durations of 300, 500, 1,000, and 3,000 ms. Participants complained that 300 ms was too brief to understand what was depicted, while results at 1,000 and 3,000 ms were very similar. Thus, for our main experiment, we chose to collect data at two durations: 500 and 3,000 ms.

Post-processing: Given the raw textual descriptions of a given image, we perform some simple text processing to form a histogram of responses (Figure 2(c)). We treat the set of responses as a "bag-of-words" [17]. Specifically, for each text description, we run a part-of-speech tagger [2], and keep only the nouns. We also discard any terms in a predefined set of 12 disallowed words, such as "abstract" and "art." Synonyms are grouped using NLTK [1], yielding a set of tokens. We then form a histogram of the tokens for the image, grouped by viewing duration. This process is performed separately for each image's responses.

Given these histograms, we measure ambiguity with two numbers: $H_{0.5}$ is the Shannon entropy of the 0.5 s viewing duration histogram, and H_3 is the entropy of the 3 s histogram. We report entropy scores in units of bits.



ΔH: 1.08 ΔH: 1.16 ΔH: 1.16 ΔH: 1.31 ΔH: 1.83

Fig. 4. Of the images in our dataset with high description entropy ($H_3 > 4.0$), the images with the lowest and highest change in entropy ($\Delta H = H_3 - H_{0.5}$). For the top row images, entropy decreased over time, and, for the bottom row, increased over time.

3 CATEGORIZING AND RANKING AMBIGUITIES

Our key assumption is that the distribution of textual descriptions for a given image and a given time duration reflect the perceptual ambiguity that a single viewer has for that image. This could arise, for example, if, when forced to describe an image, a viewer samples a single possible explanation from their posterior probability distribution. Similar processes have been hypothesized elsewhere in neuroscience [8].

Hence, the histogram of responses acts as an estimate of a typical viewer's probability distribution over image interpretations, for a given viewing duration. We can observe several types of images. In a determinate image (Figure 2 (top row)), most viewers describe the image in the same way at both viewing durations; that is, both $H_{0.5}$ and H_3 are low. In an indeterminate image (Figure 2 (bottom row)), descriptions are highly varied in both conditions.

Figure 3 shows the images with the lowest and highest entropy across descriptions. As can be seen in the figure, this metric reflects the degree of recognizability of the images. Moreover, note that none of the "AbstractFlat" images appear here. Sometimes they have high entropy (e.g., Figure 7(g)), and sometimes they have lower entropy due to a peak of color descriptions like "pink image" (e.g., Figure 7(h)). This suggests that perhaps entropy alone can identify indeterminate images. An image that is too abstract does not conjure many associations, nor does an image that is very realistic. Indeterminacy seems to produce the longest and most varied descriptions.

Figure 4 shows high-entropy images sorted by the entropy difference $H_3 - H_{0.5}$ under the condition of $H_3 > 4$. This threshold gives us images that generate more associations after 3 s viewing. When sorted by the entropy difference, we see that more complex images tend to have greater decrease in entropy over time. The complementary examples are shown in Figure 5, where the threshold is set to $H_3 < 4$. Here the images all seem to depict individual objects. Entropy appears to decrease the most where the object is odd but recognizable. Entropy appears to increase for images that seem to be complex variations on familiar objects.

Figure 6 shows a scatterplot of the entropies of images across our dataset. As shown in the plot, the image categorization that we used to build the dataset emerges in some regions of the plot; for example, recognizable images generally have lower entropy scores ($H_{0.5}, H_3 < 4$). Dichotomous images, such as Figures 7(d) and 7(e) have simple explanations at first, but diverge as viewers find two or more interpretations.

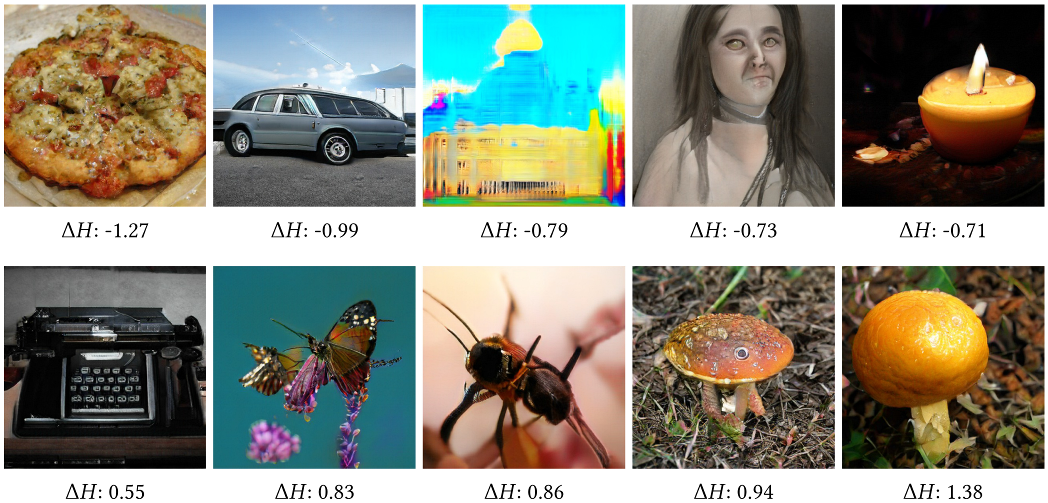


Fig. 5. Of the images in our dataset with low or medium description entropy ($H_3 < 4.0$), the images with the lowest and highest change in entropy ($\Delta H = H_3 - H_{0.5}$). Note that images with low ΔH tend to be recognizable, eventually, while large ΔH images tend to be variations on familiar concepts.

Figure 7 shows a sampling of cases with high entropy. We now discuss some of the phenomena that emerge. The first three images have high entropy ($H_3 > 4.5$), and entropy decreases over time ($H_{0.5} > H_3$). In Figure 7(a), the most-frequent terms in the 0.5 s viewing condition are "coat," "knife," and "cloth," but, in the 3 s condition, entirely new terms become prevalent, including "building," "ship," and "sails." In Figure 7(b) most descriptions in the 0.5 s condition include "face," and even more do in the 3 s condition, while several terms, like "bug," "helmet," and "monster," disappear. In Figure 7(c), "flowers" is rare in the initial condition, but becomes much more common after the longer viewing duration.

In Figures 7(d), 7(e), and 7(f), entropy increases: viewers first give consistent "first impression" descriptions, such as "buffalo," "dog," and "knife," but the descriptions become considerably more varied and diverging when viewers have spent more time viewing them.

The final images are two examples that could be seen as purely abstract art, yet viewers can occasionally perceive objects within them, though less consistently. Most tokens appear only once, as indicated by the large "[other]" category.

4 DISCUSSION

Our work is the first to attempt to go beyond the high-level impressions of ambiguity by viewers, to instead uncover and quantify how ambiguous an image is across a population of viewers. It is partly motivated by a desire to investigate whether modern computational methods can be used to address longstanding questions about subjective responses to images, particularly in light of the claims made by artists, such as Picasso, about what makes certain images aesthetically valuable.

As we show, the textual response histograms generated by our method robustly capture image properties like indeterminacy, measuring shades of ambiguity that previous studies were insensitive to.

Our preliminary study is intended to illustrate the potential value of this approach for measuring the subtle and highly subjective phenomenon of visual ambiguity. There are several limitations and directions for future development. First, we are yet to validate the indeterminacy hypothesis suggested by the art theoretical literature, that there is a positive correlation between image ambiguity and aesthetic appreciation. The tools and methods developed here provide a useful starting point from which to rigorously investigate this claim in future.

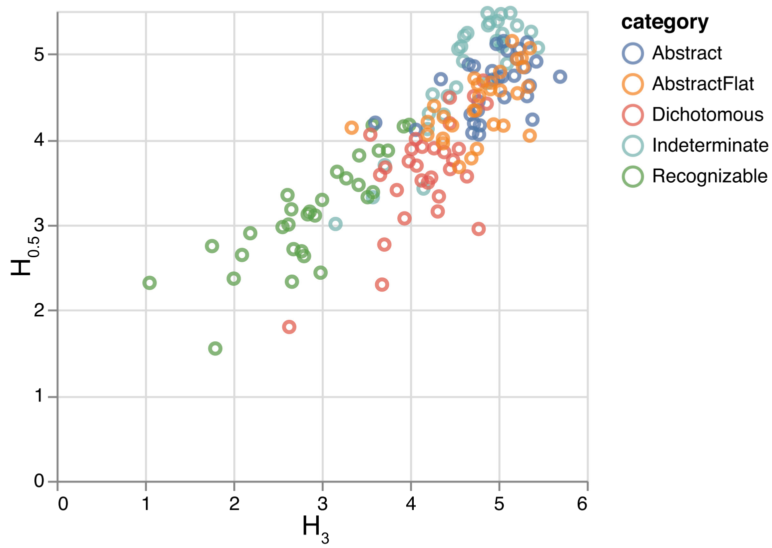


Fig. 6. Scatterplot of entropy values in our dataset. The five colors show the loose categorization used to build the dataset. Observe that the entropy values cluster some categories well, including recognizable ($H_{0.5}, H_3 < 4$) and dichotomous ($H_{0.5} < 4, H_3 > 3.5$); the indeterminate images mostly cluster toward the top. The locations of images from other figures in the article are annotated in the plot.

Second, on the basis of previous literature, we expected to find an increase in entropy over time for images with high levels of indeterminacy. We did not see this specific effect in our results (Figures 3–5), but we did find change in entropy to provide some differentiation between types of ambiguity. In fact, our results suggest that entropy alone can measure indeterminacy: If one wishes to follow Picasso's goal of producing the most associations, then an image should not be realistic, but it should also not be too abstract.

Third, our text-processing procedure is very simple. As a result, it discards some important information in the textual descriptions, e.g., it cannot distinguish between an image having diverse descriptions, because the image is confusing ("it's either a horse or a chair"), because it is dichotomous ("it's a chair made to look like a horse"), because it is complex ("it's a horse next to a chair"), or because the description is verbose ("it's a chair sitting on the ground"). Our heuristic list of disallowed words could be replaced with a more nuanced filtering. It is also unclear how best to take advantage of semantic similarity between descriptions that use distinct but related words. Our text-processing method ignored phrases that indicated difficulty or inability to respond, such as "I'm not sure but..." or "I have no idea," which occurred frequently for images in our indeterminate categories; the frequency of such phrases could be used to further inform the analysis. Application of more sophisticated text processing can yield more reliable and finer-grained insights.

Another important future step is to relate entropies to aesthetic properties. If semantic diversity and uncertainty are regarded as positive aesthetic attributes in artworks, as the art historical literature suggests, then we might expect to find a correlation between these qualities and entropy. Indeed, previous researchers have found such a positive effect [19]. As a preliminary test, we asked a separate sample of 128 crowdworkers to rate the images according to their level of "interestingness," "powerfulness," and "engagement"; however, we found that crowdworkers gave highest scores to realistic, easily-interpretable images. There may be many reasons for this finding that do not necessarily invalidate the main hypothesis of this study, including viewer expectations for which images constitute art images versus non-art images, the framing and setup of the task, and the expertise of the raters, each of which may need to be controlled for in future experiments.

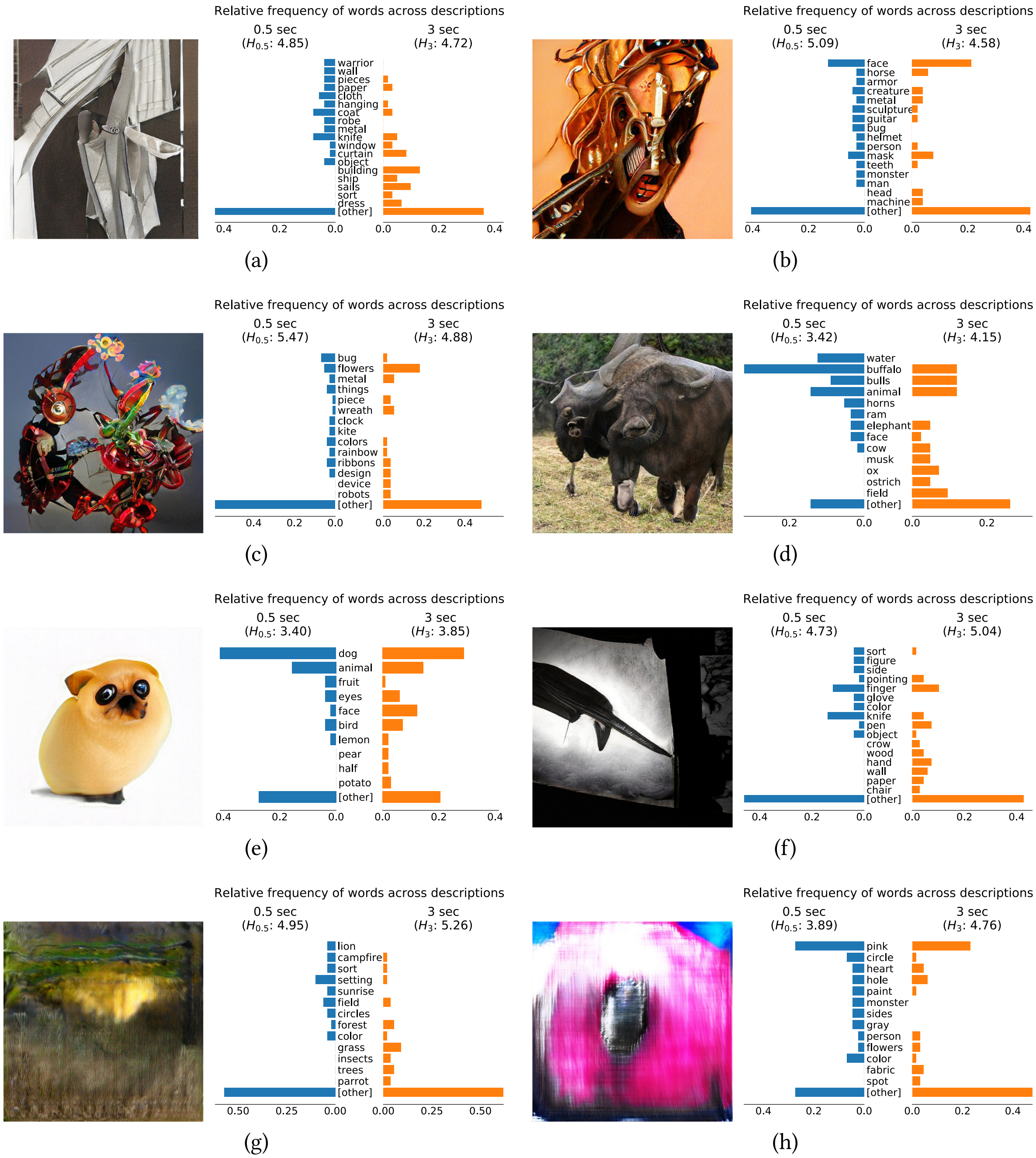


Fig. 7. More interesting examples of high-entropy images. See text for discussion.

The exposure times used in the present study were relatively short, particularly in the context of art viewing where a museum-goer might typically spend at least several tens of seconds studying an artwork to appreciate its nuances, and may return to look at it more than once [6]. Our crowdworkers often commented that they had too little time to complete their descriptions. Exposures in the order of tens of seconds or minutes may yield more complex and nuanced responses.

We thus far have only studied ambiguity of object recognition, whereas other image properties like figure-ground segmentation may also be ambiguous. However, our methodology suggests a general approach for probing perceptual uncertainty that could be combined with methods for crowdsourcing other image properties [13, 20, 32], to obtain a rich analysis of image ambiguity.

A suitably rich model of image ambiguity can open up exciting future avenues for image synthesis applications. For instance, such a model may be able to guide GANs and related image synthesis pipelines toward more interesting and unexpected creations. It may be used to automatically curate or filter image streams, or as an alternative metric of diversity to score different image synthesis techniques.

Finally, the methods being developed here could provide a potentially powerful computational framework for studying topics of interest to psychologists and neuroscientists, such as image perception, object recognition, and associative memory.

ACKNOWLEDGMENTS

We thank Aude Oliva for feedback and support for this project and for hosting X.W. as a visiting student, and Joel Simon for providing data from Artbreeder. All stimuli are public domain imagery obtained from Artbreeder, created by the following users. Figure 2: guidoheinze, kent4747; Figure 3: jakritger, caincaser, strangecircus; Figure 4: desleep, jeffgiddens, thunderdog, angrytree607, portjos, strangecircus; Figure 5: jakritger, happyemil, thelindamartinez06; Figure 7: thunderdog, desleep, spihut, telmaroza.

REFERENCES

[1] Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural Language Processing with Python.* O'Reilly Media.

[2] Tiberiu Boros, Stefan Daniel Dumitrescu, and Ruxandra Burtica. 2018. NLP-cube: End-to-end raw text processing with neural networks. In *Proceedings of the CoNLL Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies.* Association for Computational Linguistics, Brussels, Belgium, 171–179. Retrieved from http://www.aclweb.org/anthology/K18-2017.

[3] Timothy F. Brady, Talia Konkle, George A. Alvarez, and Aude Oliva. 2008. Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci. U.S.A.* 105, 38 (2008), 14325–14329.

[4] Timothy F. Brady, Talia Konkle, Jonathan Gill, Aude Oliva, and George A. Alvarez. 2013. Visual long-term memory has the same limit on fidelity as visual working memory. *Psychol. Sci.* 24, 6 (2013), 981–990.

[5] Andrew Brock, Jeff Donahue, and Karen Simonyan. 2019. Large scale GAN training for high fidelity natural image synthesis. In *Proceedings of the International Conference on Learning Representations (ICLR'19).*

[6] Claus-Christian Carbon. 2017. Art perception in the museum: How we spend time and space in art exhibitions. *i-Percept.* 8, 1 (2017), 2041669517694184.

[7] Elizabeth Cowling. 2006. *Visiting Picasso: The Notebooks and Letters of Roland Penrose.* Thames & Hudson.

[8] Nathaniel D. Daw and Aaron C. Courville. 2007. The Pigeon as particle filter. In *Proceedings of the Conference on Neural Information Processing Systems (NIPS'07).*

[9] Scott L. Fairhall and Alumit Ishai. 2008. Neural correlates of object indeterminacy in art compositions. *Conscious. Cogn.* 17 (2008), 923–932.

[10] Li Fei-Fei, Asha Iyer, Christof Koch, and Pietro Perona. 2007. What do we perceive in a glance of a real-world scene? *J. Vision* 7, 1 (01 2007), 10–10.

[11] Camilo Fosco, Anelise Newman, Pat Sukhum, Yun Bin Zhang, Nanxuan Zhao, Aude Oliva, and Zoya Bylinskii. 2020. How much time do you have? Modeling multi-duration saliency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR'20).*

[12] Dario Gamboni. 2004. *Potential Images: Ambiguity and Indeterminacy in Modern Art.* Reaktion Books.

[13] Yotam Gingold, Ariel Shamir, and Daniel Cohen-Or. 2012. Micro perceptual human computation. *ACM Trans. Graph.* 31, 5, Article 119 (Aug. 2012), 12 pages. DOI: https://doi.org/10.1145/2231816.2231817

[14] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Proceedings of the Conference on Neural Information Processing Systems (NIPS'14)*.

[15] Aaron Hertzmann. 2010. Non-photorealistic rendering and the science of art. In *Proceedings of the International Symposium on Non-Photorealistic Animation and Rendering (NPAR'10)*.

[16] Aaron Hertzmann. 2020. Visual indeterminacy in GAN art. *Leonardo* 53, 4 (2020).

[17] Thomas Hofmann. 1999. Probabilistic latent semantic indexing. In *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99)*. 50–57.

[18] Alumit Ishai, Scott L. Fairhall, and Robert Pepperell. 2007. Perception, memory and aesthetics of indeterminate art. *Brain Res. Bull.* 73, 4–6 (2007).

[19] Martina Jakesch, Helmut Leder, and Michael Forster. 2013. Image ambiguity and fluency. *PLoS ONE* 8, 9 (2013).

[20] Jan J. Koenderink, Andrea J. van Doorn, Astrid M. L. Kappers, and James T. Todd. 2001. Ambiguity and the 'Mental Eye' in pictorial relief. *Perception* 30, 4 (2001), 431–448.

[21] Claudia Muth and Claus-Christian Carbon. 2013. The aesthetic aha: On the pleasure of having insights into Gestalt. *Acta Psychologica* 144, 1 (2013), 25–30.

[22] Claudia Muth and Claus-Christian Carbon. 2016. SeIns: Semantic instability in art. *Art Percept.* 4, 1–2 (2016), 145–184.

[23] Claudia Muth, Vera Hesslinger, and Claus-Christian Carbon. 2018. Variants of Semantic Instability (SeIns) in the arts. A classification study based on experiential reports. *Psychol. Aesthet. Creativ. Arts* 12, 1 (2018).

[24] Claudia Muth, Vera M. Hesslinger, and Claus-Christian Carbon. 2015. The appeal of challenge in the perception of art: How ambiguity, solvability of ambiguity, and the opportunity for insight affect appreciation. *Psychol. Aesthet. Creativ. Arts* 9, 3 (2015).

[25] Claudia Muth, Marius H. Raab, and Claus-Christian Carbon. 2016. Semantic stability is more pleasurable in unstable episodic contexts. On the relevance of perceptual challenge in art appreciation. *Front. Hum. Neurosci.* 10 (2016), 43 pages. DOI: https://doi.org/10.3389/fnhum.2016.00043

[26] Aude Oliva. 2009. Visual scene perception. In *Encyclopedia of Perception*. SAGE Publications, Thousand Oaks, CA, 1112–1117.

[27] Robert Pepperell. 2011. Connecting art and the brain: An artist's perspective on visual indeterminacy. *Front. Hum. Neurosci.* 5 (2011), 84 pages.

[28] Robert Pepperell. 2015. Artworks as dichotomous objects: Implications for the scientific study of aesthetic experience. *Front. Hum. Neurosci.* 9 (June 2015), 295 pages.

[29] Mary C. Potter. 1975. Meaning in visual search. *Science* 187, 4180 (1975), 965–966.

[30] Mary C. Potter. 1999. Understanding sentences and scenes: The role of conceptual short-term memory. *Fleeting Memories: Cognition of Brief Visual Stimuli* (1999). Bradford, UK, 13–46.

[31] Gerhard Richter, Dietmar Elger, and Hans Ulrich Obrist. 2009. *Gerhard Richter—Text: Writing, Interviews and Letters 1961–2007*. Thames & Hudson, London.

[32] Bryan Russell, Antonio Torralba, Kevin Murphy, and William T. Freeman. 2007. LabelMe: A database and web-based tool for image annotation. In *Proceedings of the International Conference on Computer Vision (ICCV'07)*.

[33] Philippe G. Schyns and Aude Oliva. 1994. From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychol. Sci.* 5, 4 (1994), 195–200.

[34] Antonio Torralba. 2009. How many pixels make an image? *Visual Neurosci.* 26, 1 (2009), 123–131.

[35] Sander Van de Cruys and Johan Wagemans. 2011. Putting reward in art: A tentative prediction error account of visual art. *i-Percept.* 2, 9 (2011).

[36] C. Wallraven, K. Kaulard, C. Kürner, R. Pepperell, and H. Bülthoff. 2007. In the eye of the beholder - Perception of indeterminate art. In *Proceedings of the 3rd Eurographics Conference on Computational Aesthetics in Graphics, Visualization, and Imaging*. 121–128.

[37] Christian Wallraven, Kathrin Kaulard, Cora Kürner, Robert Pepperell, and Heinrich H. Bülthoff. 2007. Psychophysics for perception of (in)determinate art. In *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization*. 115–122.

Received June 2020; accepted August 2020