Beauty, the feeling

Aenne Brielmann (✉ aenne.brielmann@tuebingen.mpg.de)

Max Planck Institute for Biological Cybernetics <https://orcid.org/0000-0002-6742-2836>

Angelica Nuzzo

Graduate Center City University of New York

Denis Pelli

New York University <https://orcid.org/0000-0001-7100-5659>

Article

Keywords: beauty, feeling, pleasure

Posted Date: November 12th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-103908/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Acta Psychologica on September 1st, 2021.
See the published version at <https://doi.org/10.1016/j.actpsy.2021.103365>.

Beauty, the feeling

Aenne A. Brielmann^{1,2*}, Angelica Nuzzo³, and Denis G. Pelli^{1,4}

¹Department of Psychology, New York University, New York, NY, 10003, USA

²Department of Computational Neuroscience, Max-Planck Institute for Biological Cybernetics, Tübingen, Germany

³The Graduate Center, City University of New York, New York, NY, 10016, USA

⁴Center for Neuroscience, New York University, New York, NY, 10003, USA

*aenne.brielmann@tuebingen.mpg.de; corresponding author; lead contact

To be submitted to *Nature Human Behaviour*.

Words in abstract: 217

Figures: 2

Tables: 3

Abstract

Many philosophers and psychologists have made claims about the feelings in an experience of beauty. Here, we test how well these claims match the feelings that people report while looking at an image or listening to music that is beautiful to them, or recalling a personal experience of beauty. We conducted ten experiments (total $n = 851$) spanning three nations (US, UK, and India). Across nations and modalities, we find that top-rated beauty experiences are strongly characterized by six dimensions: intense pleasure, an impression of universality, the wish to continue the experience, perceived harmony in variety, exceeding expectation, and meaningfulness. Other frequently proposed beauty characteristics — like surprise, desire to understand, and mind wandering — are uncorrelated with feeling beauty. When explicitly asked, participants in all countries report that beauty is closely related to pleasure, found in nature rather than art, influenced by mood, and a form of communication. A remembered beautiful experience was typically active and social, like a family holiday, quite unlike the passive solitary appraisal of art emphasized in aesthetics. Our data align well with Kant ($r = 0.74$) and psychological theories that emphasize pleasure, and reject theories that emphasize information seeking. These six dimensions characterizing how individuals feel beauty emerge from the reports of hundreds of people seeing, hearing, or remembering something beautiful to them.

Introduction

Beauty has fascinated humankind for millennia. Most of the great philosophers tried to define it, including Plato (428–348 BCE; Denham, 2012), Aristotle (384–322 BCE; Halliwell, 1986), Kant (1724–1804; Nuzzo, 2005), and Hegel (1770–1831; Nuzzo, 2006). The pioneer of modern psychology, Gustav Fechner (1801—1887), took great interest in empirically testing what beauty is (Fechner, 1876). Yet, in contemporary models of aesthetic experience, beauty has mostly remained as one undefined aesthetic response among many others (e.g., Leder & Nadal, 2014; Pelowski, Markey, Forster, Gerger, & Leder, 2017).

The philosophical theories of beauty specify the contributions of many experiential dimensions to the *experience* of beauty. Many of these dimensions, including pleasure and surprise, appear in all the philosophical theories, so we can systematically score how well each theory matches contemporary feelings of beauty. In contrast, among contemporary psychological theories of beauty, each theory typically considers only the few dimensions that it supposes to be essential to experiencing beauty. These brief theories are parsimonious in order to predict beauty with just a few predictors that are not strongly correlated with each other. In contrast, the long philosophical theories aim to paint a complete picture of all the variables associated with intense beauty experiences. We here present data that speak to both approaches. To predict beauty ratings, for each participant we collect many ratings of images and music that vary in beauty. To discover correlates of beauty, we collect ratings of individual, remembered, intense beauty experiences. Together, these results reveal six experiential dimensions that characterize intense beauty.

Fechner (1876) defined beauty as an experience of pleasure. Berlyne (1971) claimed that beauty pleasure results from intermediate levels of subjective complexity or arousal. Leder and colleagues (2004) place the meaningfulness of an experience at the heart of beauty, while Vessel and colleagues (2013) identify the state of being moved as equivalent to what others call “beauty”. Diessner and colleagues (2018) equate beauty with unity in variety, i.e., a feeling that the various elements of an object are harmoniously combined into a coherent whole. (This is similar to Wittgenstein’s (1934) description of “clicking or fitting” or “harmony” in aesthetic judgement.) A growing new perspective in music science posits that passages of music are most liked, or beautiful, when the pleasure they elicit exceeds the listener’s expectation. (Salimpoor et al., 2015). While these models focus on the relation between beauty and the perceiver’s

subjective feelings, others emphasize the relation between beauty and information-seeking, broadly construed as learning, interest, or wanting to understand the experience (see also Biederman & Vessel, 2006). Armstrong and Detweiler-Bedell (2008) proposed that beauty is not only pleasurable but crucially contains an element of learning. Reber and colleagues (2004) suggested that beauty lies in the ease of processing during the experience, i.e., a feeling that the stimulus is understood. Finally, within the realm of music, Kivy (1990) claims that a piece of music needs to be interesting in order to be beautiful. It is, perhaps, due to the variety of proposed dimensions that the contemporary field of empirical aesthetics has yet to test which of these dimensions, alone or in combination, are general characteristics of beauty experiences across various kinds of object. To date, we lack a data-driven definition of what it means to experience beauty.

Therefore, we conducted ten studies in which we asked more than 800 participants to rate several dimensions, including beauty, of various experiences. To cast our net widely, we first assessed eleven of the dimensions that have been considered by prominent philosophers of aesthetics: 1) pleasure, 2) wishing to continue the experience, 3) feeling alive, 4) feeling that the experience is beautiful to everyone, 5) number of felt connections to the experience, 6) longing, 7) feeling free of desire, 8) mind wandering, 9) surprise, 10) wanting to understand the experience more, 11) feeling that the experience tells a story. Each of the seven included philosophers made statements about each of these dimensions, so we could assess whose definition fit the empirical data best. Second, we also measured people's responses to 8 additional dimensions brought forward by psychologists: 1) complexity, 2) arousal or excitement, 3) learning from the experience, 4) wanting to understand, 5) harmony in variety, 6) meaningfulness, 7) exceeding one's expectation, and 8) interest. (Our **Supplementary Material** specifies the exact wording of each question and cites the authorities who inspired it.) In addition, we also asked our participants about their explicit beliefs about beauty at the end of each experiment to probe their endorsement of seven more beauty dimensions.

While seeing, hearing, or remembering beauty, our participants rated the beauty and many other dimensions inspired by either philosophy or psychology. In this way, we identify the general characteristics of intense beauty experience across modalities. We here focus on a descriptive approach that allows us to identify experiential dimensions associated with intense

95 beauty. By uncovering which dimensions are consistently correlated with intense beauty, we
96 provide the first data-driven definition of beauty experience.
97

Results

Variables associated with beauty ratings of images and music. In our stimulus experiments, we analyzed responses by 192 participants in the USA. All together, the participants rated a total of 20 different stimuli: Experiment 1a. 5 beautiful stock-images; 5 beautiful art images; 5 neutral stock-images; Experiment 1b. 5 beautiful stock-images from Experiment 1a; the top 5 billboard musical hits of all time (details in **Methods**). We used mixed-effects linear models (Bates, et al., 2015) to assess which aspects of the experience are associated with the beauty rating. The best model was selected based on the Bayesian Information Criterion (BIC) calculated on the overall fit as well as average BIC after 10-fold cross-validation. (See **Supplementary Material** for detailed model comparisons in **Tables S23-26**, and results of all tested models in **Tables S27-29**.)

Beauty increased primarily with: perceived universality (0.23 points per point), pleasure (0.21 points per point), and a reported wish to continue the experience (0.20 points per point). Beauty was not related to surprise or the degree to which the participant felt that the stimulus told a story. Of note, the type of stimulus (music, image, or specific image kind) did not modulate these effects. Overall, this model explains 72% of the variance. **Table 1** presents the statistics of the fit. For illustration, **Figure 1A** contrasts the rating profile for stimuli rated highest (7) in beauty vs. that for those rated lower (< 7).

Table 1. The mixed-effects model that best explains rated beauty of images and music.

RANDOM EFFECTS					
	Variance	SD			
Participant	0.07	0.26			
Stimulus	0.13	0.37			
FIXED EFFECTS					
	Estimate	SE	df	t	p
Intercept	0.34	0.12	65	2.93	0.005
1. Universality	0.23	0.02	2212	12.23	<0.001
2. Pleasure	0.21	0.02	2584	11.04	<0.001
3. Wish to continue	0.20	0.02	2622	11.27	<0.001
4. Feeling alive	0.08	0.02	2645	4.30	<0.001
5. Feeling free of desire	0.07	0.02	2444	4.25	<0.001
6. Number of felt connections	0.06	0.02	2393	3.81	<0.001
7. Mind wandering	0.05	0.02	2440	3.42	0.001
8. Wanting to understand more ¹	0.04	0.02	2567	2.76	0.006
9. Longing ¹	0.04	0.02	2134	2.38	0.017
10. Telling a story	-0.01	0.01	2393	-0.73	0.466
11. Surprise	-0.02	0.01	1882	-1.13	0.260

Notes. Significant fixed effects are highlighted in bold. The model explains 72% of the variance in beauty ratings (on a scale of 1 to 7) with an RMSE of 0.96. ¹The marked dimensions did not

receive consistently high ratings for remembered beauty, which is evidence against universality of their contribution to experienced beauty.

Philosophy-based questions

Psychology-based questions

Figure 1. Average ratings for beauty experiences for all experiments. A) In USA, polar plot of average ratings on 12 philosophy-derived dimensions (12D) across all top-rated (green) vs. less-than-top-rated (red) *immediate*-beauty trials (i.e. they consider their immediate experience). B) In USA, average 12D ratings for top-rated immediate-beauty trials, separately for each stimulus type. C) Average 12D ratings on top-rated *remembered*-beauty trials (i.e. they consider their own remembered experience), separately for each independent population sample: India, UK, and two in USA. For reference, the dashed green lines represent the averages for the top-beauty-rated trials, copied from panel A. D) In USA, average ratings for all remembered-*beauty* (dark green) vs. remembered-*relief* (blue) trials. Not pictured: ratings on the two dimensions perfection and peacefulness that were not included in all studies; both peacefulness and perfection ratings were higher for beauty compared to relief ratings. E) In USA, polar plot of average ratings on 11 psychology-derived dimensions (11D) across all top-rated (green) vs. less-than-top-rated (red)

immediate-beauty trials. F) Average 11D ratings on top-rated immediate and remembered (bottom-most) trials. Shaded areas in all polar plots indicate $\pm SEM$ (not visible due to small SEM in some panels).

Variables correlated with remembered beauty. One might wonder whether a beauty experience produced in an online-test session by a previously high-rated image or popular music is likely to be strong enough to be comparable with the intense beauty experiences that philosophers wrote about. We imagine that a philosopher describing the feeling of beauty would naturally emphasize his recollection of his own most intense beauty experience. In that spirit, we asked participants in our second set of studies (Experiment 2a $n = 92$; replication Experiment 2b $n = 89$; both US American) to describe and rate a remembered intense beauty experience from their own lives. Since we intentionally collected data only for memories of intense beauty, we did not use linear models to explain beauty ratings as in Experiment 1, which would be underpowered due to the small variability in beauty ratings. Instead, we compared the ratings based on top-rated remembered- vs. immediate-beauty trials ($n = 111$ vs. 530). We here report comparisons of the five dimensions correlated with beauty in Experiment 1. Comparisons between ratings on the dimensions not correlated with intense beauty can be found in the **Supplementary Material**.

According to Kolmogorov-Smirnov tests, ratings of pleasure, wishing to continue the experience, and mind wandering were not differently distributed for immediate- vs. remembered-beauty trials, both $p \geq 0.448$. Yet, participants rated remembered-beauty higher in terms of feeling alive, feeling free of desire, and the number of felt connections, all $0.41 \leq d \leq 0.63$. In contrast, they rated remembered beautiful experiences as less universally beautiful, wanted to understand them more to a lesser extent, and reported less longing compared to immediate beauty experiences, all $0.32 \leq d \leq 0.68$. Nonetheless, ratings on all seven of the nine dimensions correlated with beauty ratings in Experiments 1a and 1b were also high for remembered beauty, all means ≥ 5.43 on a 1-7 scale. The exceptions were ratings of wanting to understand the experience more and feeling alive, both means $\leq .76$, rendering these dimensions less likely to be essential constituents of beauty experiences. This is unsurprising given the weak association with beauty revealed by the general linear model.

Figures 1A vs. C illustrate the similarity of average ratings in Experiments 1 vs. 2: immediate vs. remembered beauty. (To facilitate this comparison, the top-beauty data from **Fig. 1A** appear as dashed lines in the USA panels of **Fig. 1C**.)

Conservation across cultures. In our third set of studies, we asked how well our previous findings generalize across countries and cultures. We therefore compared the beauty memory ratings from our initial US samples to those from the UK ($n = 78$) and India ($n = 52$). (Our questionnaire is in English, so we confined ourselves to English-speaking countries.) There were differences in the average rating pattern between countries, $F(2,304) = 3.50$, $p < 0.001$. Because average beauty ratings differed, too, we restricted the comparison between countries to top-rated remembered-beauty trials ($n = 25$ for India; $n = 56$ for UK; $n = 111$ for USA). Of the seven dimensions consistently associated with high beauty ratings, only two differed across countries. The impression of universality was higher in India, $M = 6.36$, than in the USA, $M = 5.43$, $p = 0.018$, $d = 0.62$, and reported mind wandering was lower in the USA, $M = 5.37$, than the UK, $M = 6.05$, and India, $M = 6.32$, both $p \leq 0.024$, both $d \geq 0.41$. Notably, longing ratings were much higher in India, $M = 5.44$, than any other country, both $M \leq 3.67$, both $p \leq 0.001$, potentially suggesting a greater import of this dimension in India than other English-speaking countries. In sum, the general pattern of ratings correlated with beauty was nearly identical across cultures.

Uniqueness of the beauty-rating profile: a comparison to relief. So far, we have described the characteristics of a beauty experience, but we have not addressed which of these characteristics might distinguish beauty from other positive experiences. To do so, we asked an independent sample of US Americans ($n = 90$) to recall and then describe and rate a personal experience of intense relief. We compare beauty to relief, here, because it is a strongly positive emotion, the memory of which is not usually also deemed beautiful (in contrast to, e.g., joy, see **Supplementary Material**). The ratings for remembered relief are very different from those for remembered beauty (see **Figure 1D**), $F(1,395) = 0.53$, $p < 0.001$. Follow-up t-tests indicate that the two concepts do not differ in perceived universality, $p = 0.381$, or surprise, $p = 0.471$. However, remembered relief received lower ratings on all 12 remaining dimensions, all $p \leq 0.001$, all $d \geq 0.39$. Thus, among the dimensions correlated with intense beauty, all but one — universality — are *uniquely* correlated with beauty, not just any positive memory.

Comparison between responses and philosophers' theories. Of the seven philosophers considered here, Kant has the most matches to the data: 10 out of 11 assessed dimensions. (See **Supplementary Table S30** for correlations between data and all 7 included philosophers.) Kant's theory also states our participants' belief that beauty is found in nature rather than art and

that both the object as well as its story contribute to beauty. Yet, contrary to Kant's theory, higher surprise is not associated with more intense beauty. However, as we will show below, there is a kind of surprise, i.e., the surprise of something exceeding one's expectation, that is indeed linked to beauty.

Comparison between responses and psychological theories. Our first series of studies focused on questions based on philosophical theories of beauty because they all make statements regarding a wide range of characteristics. However, we also wanted to assess how well modern psychological theories of beauty reflect people's beauty experiences. We therefore ran modified versions of Experiments 1a, 1b, and 2b with questions gleaned from psychological theories about beauty (see **Table 2** for a complete list).

We again assessed which of the 11 tested dimensions were associated with beauty ratings using mixed-effects linear models for data of those participants who rated images and/or music (total $N = 186$). As in Experiment 1, we tested several models. Again, we created two sets of models: In one set, complexity and excitement were added as linear terms, like the other dimensions. In the other set – based on Berlyne's claims (1971) – they were added as squared terms. According to the average BIC based on 10-fold cross-validation, a model that includes an interaction with stimulus modality explains beauty ratings best, explaining 70% of the variance (see **Supplementary Tables S31-32** for detailed model comparisons). For all stimuli, beauty increases with pleasure (0.28 points per pleasure point for images, 0.07 for music), feeling moved (0.18 beauty points per point), perceived harmony among variety (0.11 beauty-points per point), feeling that the experience exceeded expectation (0.11 points), and meaningfulness (0.08 points for images, 0.21 points for music). In addition, beauty ratings were inversely correlated with squared complexity ratings. Of note, interest was correlated with beauty ratings for images, but not for music. The detailed parameters of the model are listed in **Table 2**.

Next, we again compared ratings on all dimensions for immediate experiences to those for the remembered beauty experiences of an independent sample ($N = 72$). Ratings of pleasure, feeling that the experience exceeded expectation, and harmony in variety did not differ between stimulus- and memory-related beauty experiences with top (7) ratings, all $p \geq 0.265$. Memories were, however, rated slightly more moving, meaningful, and relatively less complex, all $p \leq 0.040$. Of note, all these differences were small, all $|d| \leq 0.45$, and all means were in the range expected based on the linear model (see **Table 2**). Thus, as was the case for the dimensions

collected from philosophical theories, the psychology-theory related characteristics of beautiful experiences did not differ much between modalities.

Table 2. The mixed-effects model that best explains rated beauty of images and music based on questions from psychological theories.

RANDOM EFFECTS					
	Variance	<i>SD</i>			
Participant	0.06	0.25			
Stimulus	0.05	0.21			
FIXED EFFECTS					
	Estimate	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
(Intercept)	0.55	0.13	287	4.09	<0.001
Music	-0.49	0.27	312	-1.86	0.063
Pleasure	0.28	0.02	2221	12.00	<0.001
Moved	0.18	0.02	2229	8.09	<0.001
Exceeded expectation	0.11	0.02	2148	5.05	<0.001
Harmony in variety	0.11	0.02	2219	5.44	<0.001
Meaningful	0.08	0.02	2227	3.34	0.001
Complex (squared)	-0.02	0.01	2192	-2.58	0.010
Interesting	0.17	0.02	2225	7.61	<0.001
Exciting (squared)	-0.01	0.01	2215	-0.81	0.418
Learning from the experience	-0.02	0.02	1482	-1.13	0.257
Understandable	0.03	0.02	2144	1.53	0.127
Music × pleasure	-0.21	0.06	2204	-3.64	<0.001
Music × moved	-0.03	0.06	2222	-0.57	0.568
Music × exceeded expectation	-0.02	0.05	2212	-0.47	0.638
Music × harmony in variety	-0.07	0.05	2236	-1.37	0.170
Music × meaningful	0.14	0.05	2228	2.52	0.012
Music × complex (squared)	0.01	0.02	2226	0.29	0.773
Music × interesting	-0.15	0.06	2227	-2.76	0.006
Music × exciting (squared)	0.00	0.02	2223	0.07	0.941
Music × learning from the experience	0.29	0.04	2217	6.53	<0.001
Music × understandable	0.08	0.05	2230	1.59	0.113

Notes. Significant fixed effects are highlighted in bold. The model explained 70 % of the variance in beauty ratings (on scale of 1 to 7) with an RMSE of 0.94.

Comparison between responses and psychologists' definitions. The ratings for images, music, and memories reveal a link to beauty for seven out of the eleven here-considered characteristics that psychologists claim are linked to beauty. In all of our surveys, intense beauty was associated with intense pleasure, as claimed by Fechner (1876), the first psychologist to do so. A strong link between beauty and being moved was also evident, as reported by Vessel and colleagues (2013). The notion that a positive prediction error contributes to beauty (Salimpoor, et al., 2015) was also confirmed. Harmony in variety, the central beauty criterion in Diessner and colleague's (2018) theory was associated with beauty, too, and so was meaningfulness (see Leder et al., 2004). We found mixed results regarding Berlyne's (1971) claims. While our results indeed suggest an inverted-u-shaped relation between subjective complexity and beauty, where

medium complexity levels are associated with most intense beauty, we did not find such a relationship between beauty and arousal (assessed as excitement). Our music results reject Kivy's (1990) claim that musical beauty is linked to interest.

Information-seeking — i.e. learning, interest, and wanting to understand the experience— was not systematically linked to beauty in our studies, contrary to previous claims (Biederman & Vessel, 2006). Our results reject Reber and colleague's (2004) claim that ease of processing— i.e., understanding—of the experience is tied to its beauty. The other cognitive dimension we included, learning, was also not associated with beauty, contrary to claims by Armstrong and Detweiler-Bedell's (2008). Taken together, these negative findings indicate that information-seeking is not important for experiencing beauty.

Meaningful. In his essay “The abuse of beauty”, the philosopher and art critic Arthur Danto (2002) notes that until World War I and the dada movement, it was generally accepted that beauty was central to the definition of art. Danto notes that the dada and the subsequent postmodern movements “disconnect[ed] beauty from art”. Art today is much more general than just beauty. Danto says that beauty is merely one of many attributes that art can have and that the only necessary one is meaning. However, Danto's strict dichotomy between beauty and meaning is undermined by finding that beauty is associated with meaningfulness in our participants' reports.

Text analyses. To achieve a characterization of beauty experiences independent of our rating scales, we analyzed the beauty memory descriptions that our participants wrote down before rating them (available with all data at https://github.com/aenneb/characterizing_beauty). **Table 3** lists the top ten most frequently used words (excluding stop-words) for each experiment (extracted using NLTK; Loper, & Bird, 2002). Going beyond the count of word frequencies, we used the *empath client* (Bates, Maechler, Bolker, & Walker, 2015) to analyze which lexical categories are most represented in our beauty descriptions compared to a standard text corpus. We find that the top ten lexical categories in the beauty memory descriptions were, in order: beauty, attractive, feminine, weather, children, love, beach, vacation, positive emotion, and party. These stand in stark contrast to most of the top themes emerging in relief memories, i.e., negative emotion, contentment, joy, traveling, party, family, driving, home, listen, pain. Thus, the remembered experiences were typically active and social, like a family holiday, unlike the passive solitary appraisal of art emphasized in aesthetics.

The **Supplementary Material** contains detailed text analyses for each experiment. Writings in aesthetics emphasize passive solitary appraisal of art, but the beautiful experience recalled by our participants was more typically active and social, like a family holiday.

Table 3. Top-ten most frequently appearing words across beauty memory descriptions per experiment

Frequency rank	Exp 2a (USA)	Exp 2b (USA)	Exp 3a (UK)	Exp 3b (India)	Exp6 (USA)	All
1	Beautiful	Beautiful	Day	Beautiful	Beautiful	Beautiful
2	Experience	Beauty	Beautiful	Felt	Time	Day
3	Could	Felt	Time	Feel	Day	Time
4	Like	Time	Could	Day	Beauty	Beauty
5	Beauty	Day	Like	Beauty	Remember	Like
6	See	Like	See	Time	One	Felt
7	One	Went	Felt	Like	See	One
8	Time	Nature	Smell	Experience	Went	See
9	Went	One	Beauty	Life	Mountains	Went
10	Remember	First	Sun	Place	Feel	Could

Notes. Counts for each word per experiment are provided in the Supplementary Material.

People’s explicit beliefs about beauty. Participants in all experiments answered the same six questions about their general beliefs about beauty at the end of the experiment (total $n = 851$). As illustrated in **Figure 2**, participants in all countries did endorse the statement that pleasure and beauty are closely related, that sharing beauty is a form of communication, and that mood influences beauty, $Md = 6$ each on a 1 (not at all) to 7 (very much) scale. They also perceived beauty to lie in nature more than in art, $Md = 6$ on a 1 (art) to 7 (nature) scale, and both within the object itself as well as in the story it tells, $Md = 4$. The only question that divided people was whether a universally beautiful object exists: About half of our participants (56.29%) agreed, the remainder did not. We asked those participants to tell us of such an object. Their answers fell into a few categories. The great majority (78% of 479) named an element of nature, mostly flowers (22%), the sky, sun, or related phenomena (22%). Of the much rarer non-nature-related answers, the most common ones referred to valuable objects (6%), mostly diamonds, or artworks (broadly construed: 6%). Participants who answered questions based on psychological theories ($N = 258$) were also asked whether all beauty experiences are fundamentally the same, since this statement is the central claim emerging from a series of fMRI studies (Ishizu & Zeki, 2011). About half of these participants agreed with the statement (56%). This answer was somewhat related to people’s belief in a universally beautiful object. 64% of participants who said there is a universally beautiful object also said that all beauty experiences are the same,

306 whereas only 44% of those who did not believe in a universally beautiful object said that all
307 beauty experiences are the same.
308

309
310 Figure 2. Histograms of ratings in response to the seven questions on participants' general
311 beliefs about beauty posed at the end of each experiment. Each solid bar indicates the

proportion per experiment, differentiated by shades of gray. Immediate beauty USA (2) and remembered beauty USA (3) refers to the people who rated psychology-based questions. Open bars indicate the overall distribution of ratings across all experiments.

Implications for the science of beauty. Our results provide an empirical characterization of the beauty experience. They will inform current theories of aesthetic appreciation (e.g., Leder & Nadal, 2014; Pelowski et al., 2017) and provide a first broad test of philosophy- and psychology-based theories of beauty. Our findings complement previous efforts that contrasted people's theoretical conceptions of beauty to other aesthetic evaluations (Menninghaus, Wagner, Kegel, Knoop, & Schlotz, 2019; see **Supplementary Material** for a quantitative comparison between their and our data).

Our findings are in line with the notion that beauty is a positive emotion, i.e., strongly correlated with pleasure (Armstrong & Detweiler-Bedell, 2008; Fechner, 1876) and being moved (Vessel et al., 2013). We did not find a correlation between intense beauty and information-seeking here, i.e., wanting to understand the experience, learning, or interest, despite the prevalence of these notions in several psychological theories (Armstrong & Detweiler-Bedell, 2008; Kivy, 1990; Reber, Schwarz, & Winkielman, 2004). We did, however, find that several other features mentioned in contemporary literature were indeed correlated with beauty, such as the feeling that the experience exceeded expectation (Salimpoor et al., 2015), harmoniously combined various elements (Diessner et al., 2018), and meaningfulness (Leder et al., 2004).

The descriptions of beautiful memories our participants provided as well as their explicit beliefs about beauty call the focus on art in models of aesthetic appreciation into question. Participants endorsed the statement that beauty lies in nature, rather than art, and a typical memory of an intense beauty experience was a family vacation on the beach or the mountains, rather than a museum visit. In addition, most people who did believe in a universally beautiful object thought it to be an element of nature, like flowers or the sky, and only a small minority named artworks of any kind. Thus, research on aesthetic appreciation that wants to claim relevance for everyday beauty experiences should include such social and nature-related experiences.

Our current study documents what feelings are correlated with beauty. This list of beauty-associated characteristics offers a basis for developing a predictive model. The development of a model that can predict the beauty intensity of a given experience with as few predictors as possible would be a big step in explaining beauty.

Conclusion. The experience of beauty is characterized by: intense pleasure, an impression of universality, wanting to continue the experience, perceived harmony in variety, exceeded expectation, and meaningfulness. This holds true for images, music, and memories across three culturally-diverse English-speaking countries. The remembered experiences were typically active and social, like a family holiday, unlike the passive solitary appraisal of art emphasized in aesthetics. Among seven renowned philosophers of aesthetics, Kant's definition best matches the results of our survey of the feeling of beauty. Our results are in line with several psychological theories of beauty but not with those that emphasize information seeking. This broad new characterization of the feeling of beauty comes from the answers of 851 people to questions about their top-rated beauty experiences, both immediate and recalled.

Author contributions AAB co-designed the study, collected and analyzed the data, wrote the first manuscript draft, and created figures and tables. AN, as a philosopher, prepared the rating questions and estimated how the philosophers (except Danto) would have responded. DGP co-designed the study and recruited AN. All authors discussed the results and wrote the paper.

Acknowledgements We thank David Brainard for suggesting the memory experiment. Thanks to Amy Belfi and Elena Bai for providing the coding of beauty memories in terms of internal and external focus. Thanks to Alexandra R. Lang for her contributions to the planning stage of the first experiments. Thanks to Lily Liu for estimating the correlation between 11 parameters and beauty according to Danto. Thanks to David Hobbs for his contributions to the initial exploration of the beauty memory descriptions. Thanks to Gary Tomlinson for telling us about surprise in music. Thanks to Ashley Feng and Laura Suciú for helpful comments. This work was supported by NIH Core Grant P30 EY013079.

Competing financial interests. All authors declare no competing financial interests.

Contact. Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Aenne Brielmann (aenne.brielmann@nyu.edu).

Data and software availability. All data and main analyses files for this article are accessible on GitHub: https://github.com/aenneb/characterizing_beauty.

References

- Armstrong, T., & Detweiler-Bedell, B. (2008). Beauty as an emotion: The exhilarating prospect of mastering a challenging world. *Review of general psychology*, 12(4), 305-329.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. [doi:10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Bartoń, K. (2019). MuMIn: Multi-Model Inference. R package version 1.43.6. <https://CRAN.R-project.org/package=MuMIn>
- Biederman, I., & Vessel, E. A. (2006). Perceptual pleasure and the brain: A novel theory explains why the brain craves information and seeks it through the senses. *American scientist*, 94(3), 247-253.
- Brielmann, A. A., & Pelli, D. (2019). Intense beauty requires intense pleasure. *Frontiers in psychology*, 10, 2420.
- Danto, A. C. (2002). The Abuse of Beauty, *Daedalus*, Vol. 131, No. 4, On Beauty (Fall, 2002), pp. 35-56. <http://www.jstor.org/stable/20027805>
- Denham, A. E. (Ed.) (2012). *Plato on Art and Beauty*. New York, NY: Palgrave MacMillan.
- Diessner, R., Pohling, R., Stacy, S., & Güsewell, A. (2018). Trait appreciation of beauty: A story of love, transcendence, and inquiry. *Review of General Psychology*, 22(4), 377-397.
- Fast, E., Chen, B., & Bernstein, M. S. (2016, May). Empath: Understanding topic signals in large-scale text. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4647-4657). ACM.
- Halliwell, S. (1986). *Aristotle's Poetics*. Chapel Hill, NC: University of North Carolina Press.
- Ishizu, T., & Zeki, S. (2011). Toward a brain-based theory of beauty. *PloS one*, 6(7).
- Kivy, P. (1990). *Music alone: Philosophical reflections on the purely musical experience*. Cornell University Press.
- Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior research methods*, 49(2), 457-470.
- Kuznetsova A., Brockhoff P. B., Christensen R. H. B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, 82(13), 1–26. [doi: 10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13).
- Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode—Developments and challenges in empirical aesthetics. *British Journal of Psychology*, 105(4), 443-464.
- Loper, E., & Bird, S. (2002). NLTK: the natural language toolkit. *arXiv preprint cs/0205028*.
- Menninghaus, W., Wagner, V., Kegel, V., Knoop, C. A., & Schlotz, W. (2019). Beauty, elegance, grace, and sexiness compared. *PLOS ONE*, 14(6), e0218728.
- Nuzzo, A. (2005). *Kant and the Unity of Reason*. West Lafayette, IN: Purdue University Press.
- Nuzzo, A. (2006). Hegel's 'Aesthetics' as Theory of Absolute Spirit. *Internationales Jahrbuch des Deutschen Idealismus/International Yearbook of German Idealism*, 4, 291-310
- Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me... delight my eyes and brain: The Vienna integrated model of top-down and bottom-up processes in art perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, 21, 80-125.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?. *Personality and social psychology review*, 8(4), 364-382.
- Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: how musical sounds become rewarding. *Trends in cognitive sciences*, 19(2), 86-91.
- Vessel, Edward A., G. Gabrielle Starr, and Nava Rubin. "Art reaches within: aesthetic experience, the self and the default mode network." *Frontiers in Neuroscience* 7 (2013): 258.
- Vessel, E. A., Maurer, N., Denker, A. H., & Starr, G. G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. *Cognition*, 179, 121-131.
- Wittgenstein, L. (1938). *Lectures & Conversations on Aesthetics, Psychology and Religious Belief*. University of California Press, Berkeley and Los Angeles.

Methods

For each experiment, we recruited 100 or more participants via Amazon mechanical Turk (mTurk). Inclusion criteria were: mTurk workers with an approval rate of at least 90%, who have completed at least 50 HITs, and reside in the country targeted in the particular experiment. All participants consented to participate according to a consent form approved by the NYU UCAIHS (university committee on activities involving human subjects; IRB-FY2016-404) by checking a box in the online form. Participants were reimbursed \$15 per hour; the payment depended on the duration of the experiment. All data were collected between April 2019 and May 2020.

General methods.

Procedures. Participants rated images (Experiment 1a), images and music (Experiment 1b) or memories (Experiment 2-4) on the same 12 dimensions: beauty, pleasure, surprise, wanting to experience the stimulus longer, feeling free of desire, feeling alive, wanting to understand the experience more, mind wandering, number of connections felt with the experience, how far the experience tells a story, how far the experience is beautiful (or relief in Experiment 4) for everyone, and longing. Participants in Experiments 2-4 rated experiences on a further 2 dimensions. In Experiment 2b, an additional set of four questions was added as well as a question about whether they had the experience alone. All questions are listed in the **Supplementary Material**. All ratings were given on a scale from "not at all" (1) to "very much" (7), except for the scale asking about the number of connections with the experience ("None" (1) to "Many" (7)). At the end of each experiment, participants also answered six questions about their general beliefs about beauty, their age and gender. Participants in Experiments 1a and 2b also provided additional demographic information. These additional questions were removed from other experiments because none of these additional variables correlated with the ratings.

Analyses. After initial inspection, data was analyzed with R version 3.5.3 or higher, python version 3.6.8 or higher, and MATLAB version R2018b.

Experiment 1a: Rating images of varying beauty.

Participants. Of the 100 recruited participants, 99 completed the survey. Of these, 66 were male, 31 female, and 2 did not disclose their gender. Their ages ranged from 20 to 72 with a mean age of 34.8 ($SD = 9.9$). Most had earned a college degree ($n = 50$), or at least some college

education ($n = 24$). The remaining participants either had a high school ($n = 17$) or graduate degree ($n = 8$). Most participants had neither any formal art ($n = 83$) nor any philosophy education ($n = 87$). Few attended some art ($n = 13$), or philosophy courses ($n = 11$). Only three participants had a degree in art or art history, and only one had a degree in philosophy. Most participants' household income ranged between \$50,000 and \$70,000 ($n = 42$), with 17 falling below that range, and 40 above. Most ($n = 73$) participants were white, seven identified as American Indian and multiracial each, five identified as black and Hispanic each, and one identified as South Asian and other each. On average, participants identified as more liberal than conservative on a 1 (very liberal) to 8 (very conservative) scale, $M = 3.3$, $SD = 1.9$.

Stimuli. We selected the five most beautiful image from the open affective standardized image set (OASIS; Kurd, Lozano, & Banaji, 2017) based on a previous study conducted by our lab (Brielmann & Pelli, 2019), as well as three images with median beauty ratings. The OASIS consists of a diverse set of stock images. We will therefore refer to these stimuli as *beautiful stock-images* (B) and *neutral stock-images* (N), respectively. In addition, we selected the five most beautiful art images of a diverse set of paintings previously used by Belfi and colleagues (2019). We will refer to these images as *beautiful art-images*.

Procedures. On each trial, participants saw one image with a rating scale below it. The image and rating scale were displayed until the participant did the rating and clicked the “next” button. Thus, each trial got one rating for one stimulus. Four beautiful art- and stock-images and two neutral stock-images were rated on each of 12 dimensions. One image of each category was rated twice on each dimension, the remaining stimuli were rated once. The order of images and ratings was randomized for each participant.

Analyses. For all analyses involving demographics, we binarized the art and philosophy education demographic into “has” (either some courses or degree) vs. “has not”. We fit linear mixed-effects models with the R package *lme4* (Bates, Maechler, Bolker, & Walker, 2015) and obtained further statistics with the *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) and *MuMIn* packages (Bartón, 2019). The linear models predicted beauty ratings. We successively evaluated linear mixed-effects models, starting with the simplest (only including random effects of either stimulus or participant), then adding fixed effects of all remaining eleven ratings, and lastly exploring the interaction of demographic variables with the linear combination of the eleven ratings. We used the built-in ANOVA comparison of models to find the best fitting model

considering the number of parameters. Differences between continuous variables were tested with two-tailed t-tests, those between ordinal variables with Wilcoxon-rank-sum tests and those between proportions with the built-in *prop.test* function in R. We used MATLAB R2018b to run cluster analyses.

Experiment 1b: Rating beautiful images and music

Participants. Of the 100 recruited participants, 99 completed the survey. We excluded an additional six participants due to $\geq 5\%$ wrong responses on the question asking about the stimulus type (see **Procedures** below). We thus analyzed data from 93 participants. Of these, 51 were male, and 42 female. Their ages ranged from 18 to 64 with a mean age of 36.8 ($SD = 10.8$).

Stimuli. Participants saw the same five most beautiful images from the OASIS (Kurdi Lozano, & Banaji, 2017) as used in Experiment 1a. In addition, they listened to the “greatest-hot-100-singles of all time” according to the music billboard charts (<https://www.billboard.com/charts/greatest-hot-100-singles>). Both images and music were chosen so as to maximize the beauty rating.

Procedures. The procedures were identical to Experiment 1a except for the changes necessary to ensure that participants listened to the music. To do so, the main experiment was preceded by explicit instructions to turn on speakers or headphones and a sound check question. In 50% of the trials in this experiment, one of the five songs started to play from the start. After each trial, we asked participants whether they saw an image or listened to a song with an additional open-ended “other” option. Trials in which “other” was selected were excluded from the analyses.

Analyses. We used the same linear mixed model analyses as in Experiment 1a.

Experiment 2a: Rating remembered beauty

Participants. Of the 100 recruited participants, 99 completed the survey. Based on the written memory descriptions provided, we excluded seven participants (4 men, 3 women) due to apparent non-compliance. Of the remaining 92 participants, 71 were male, 21 female. Their ages ranged from 19 to 70 with a mean of 34.3 years ($SD = 10.2$).

Procedures. We told participants to “Please think back to an experience during which you felt intense beauty. Picture the experience. Remember as many details as you can: what you

saw, heard, smelled, and felt. Let the memory linger for a minute.” A timer counted down one minute in 10-second intervals. Participants were only able to continue the survey after a minimum of one minute had elapsed. Next, we asked them to provide a written description of their experience. On the next page, we asked them how long ago the experience had occurred, and let them rate the experience on the 12 dimensions listed in **General Methods** plus two dimensions that we added based on further discussion of the results of Experiment 1, namely perfection and peacefulness.

Analyses. We compared ratings for top-rated immediate- and remembered-beauty trials (i.e., trials with beauty ratings = 7). For the data of Experiment 1, this means that we included 1 to 18 trials of 71 participants, a total of 356 trials. Fifty-three of the 90 participants in Experiment 2 were included. We used the python package *NLTK* (Loper & Bird, 2002) and *empath-client* (<https://github.com/Ejhfast/empath-client>; Fast, Chen, & Bernstein, 2016) to analyze written memory descriptions.

Experiment 2b: replication and extension of Experiment 2a

Participants. One hundred and one participants completed the survey. Based on the written memory descriptions provided, we excluded 12 participants due to apparent non-compliance. Of the remaining 89 participants, 57 were male, 32 female. Their ages ranged from 19 to 64 with a mean of 34.9 years ($SD = 11.0$). Similar to our first experiment, we collected more extensive demographic information regarding education as well as religion, which is available in the **Supplementary Material**.

Procedures & analyses. The main procedures and analyses were identical to Experiment 2a with two exceptions.

One, we added a perceptual task to assess response biases. Participants were shown two gray circles for one second and asked to rate which one was bigger and by how much. Importantly, the latter rating was given on the same scale from “Not at all” (1) to “Very much” (7) scale than all other ratings. The right circle’s diameter was 90% of the left circle’s. With this rating, we assessed potential response biases by correlating the circle-size comparison rating with ratings on all other dimensions using Pearson’s correlations.

Two, we included six new questions after reviewing the data from Experiments 3-6: 1) “Did this experience give you a new perspective on other experiences?”, 2) “How strong do you

think this experience would be if you had it again?”, 3) “Would sharing this experience with friends make it better?”, 4) “Did this experience give you a new perspective on yourself?”, 5) “Were you alone when you had this experience?”, 6a) if “yes” to 5: “Did you wish you could share this experience with others?”, 6b) if “no” to 6: “Did the experience make you feel more connected with the people you were with?”.

Experiment 3a: Remembered beauty in the UK

Participants. We initially aimed to recruit 100 participants from the UK via amazon mechanical Turk. Of the 100 recruited participants, 99 completed the survey. Based on the written memory descriptions provided, we excluded 21 participants due to apparent non-compliance. Of the remaining 78 participants, 51 were male, 27 female. Their ages ranged from 19 to 70 with a mean of 34.3 years ($SD = 10.2$).

Procedures & analyses. All procedures and analyses were identical to Experiment 2a.

Experiment 3b: Remembered beauty in India

Participants. Even though we requested only 150 participants, 154 completed the survey. Based on the written memory descriptions provided, we excluded 102 participants due to apparent non-compliance: including several duplicate answers, copy-pasted responses unrelated to the task at hand, and one-word beauty memory descriptions. Of the remaining 52 participants, 43 were male, 9 female. Their ages ranged from 22 to 40 with a mean of 27.9 years ($SD = 4.4$).

Procedures & analyses. The main procedures and analyses were identical to Experiment 2b. The exception was that we decreased the size of the smaller circle for the perceptual task to 80% the diameter of the larger one.

Experiment 4: Remembered relief (USA)

Participants. We recruited 100 participants from the USA via amazon mechanical Turk. All completed the survey. Based on the written memory descriptions provided, we excluded 10 participants due to apparent non-compliance. Of the remaining 90 participants, 57 were male, 33 female. Their ages ranged from 20 to 69 with a mean of 35.8 years ($SD = 12.1$).

Procedures & analyses. All procedures and analyses were identical to Experiment 2a. To compare 14D ratings of remembered beauty and relief, we first tested whether an overall difference exists with a MANOVA, followed up with separate two-sided t-tests for each rating.

Experiments 5-6: Rating stimuli on dimensions derived from psychology literature

After establishing which dimensions considered by philosophers were correlated with people's beauty experiences, we turned to contemporary theories of beauty to test which of their suggested features correlate with beauty. To do so, we replicated Experiments 1a, 1b, and 2b with a different set of questions derived from the main psychological theories that make statements about the experience of beauty (see **Supplementary Material** for details and references). All stimuli and instructions were kept the same. We list the most important information about the participants below. More extensive demographic information is available in the **Supplementary Material**.

Experiment 5a – rating images. One hundred and one participants completed the survey. Of these, 63 were male, and 38 female. Their ages ranged from 20 to 73 with a mean age of 37.8 ($SD = 11.9$).

Experiment 5b – rating images and music. One hundred participants completed the survey. We excluded fifteen participants due to $\geq 5\%$ wrong responses on the question asking about the stimulus type (see Experiment 2b above). We thus analyzed data from 85 participants. Of these, 48 were male, and 37 female. Their ages ranged from 22 to 72 with a mean age of 37.8 ($SD = 11.5$).

Experiment 6 – rating memories. Of the 100 recruited participants, 99 completed the survey. Based on the written memory descriptions provided, we excluded 27 participants due to apparent non-compliance. Of the remaining 72 participants, 34 were male, 36 female, one identified with another gender, and one preferred not to answer the question. Their ages ranged from 22 to 69 with a mean of 39.5 years ($SD = 12.5$).

Figures

Philosophy-based questions

A Immediate beauty. Top vs. not. (USA)

B Immediate beauty. Kind of stimulus. (USA)

C Remembered beauty

D Remembered beauty vs. relief (USA)

Psychology-based questions

E Immediate beauty. Top vs. not. (USA)

F Kind of experience. (USA)

Figure 1

Average ratings for beauty experiences for all experiments. A) In USA, polar plot of average ratings on 12 philosophy-derived dimensions (12D) across all top-rated (green) vs. less-than-top-rated (red) immediate-beauty trials (i.e. they consider their immediate experience). B) In USA, average 12D ratings for top-rated

immediate-beauty trials, separately for each stimulus type. C) Average 12D ratings on top-rated remembered-beauty trials (i.e. they consider their own remembered experience), separately for each independent population sample: India, UK, and two in USA. For reference, the dashed green lines represent the averages for the top-beauty-rated trials, copied from panel A. D) In USA, average ratings for all remembered-beauty (dark green) vs. remembered-relief (blue) trials. Not pictured: ratings on the two dimensions perfection and peacefulness that were not included in all studies; both peacefulness and perfection ratings were higher for beauty compared to relief ratings. E) In USA, polar plot of average ratings on 11 psychology-derived dimensions (11D) across all top-rated (green) vs. less-than-top-rated (red) immediate-beauty trials. F) Average 11D ratings on top-rated immediate and remembered (bottom-most) trials. Shaded areas in all polar plots indicate $\pm$ SEM (not visible due to small SEM in some panels).

Figure 2

Histograms of ratings in response to the seven questions on participants' general beliefs about beauty posed at the end of each experiment. Each solid bar indicates the proportion per experiment, differentiated by shades of gray. Immediate beauty USA (2) and remembered beauty USA (3) refers to the people who rated psychology-based questions. Open bars indicate the overall distribution of ratings across all experiments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [BeautyTheFeelingSUPPLEMENT.pdf](#)