"Taste typicality" is a foundational and multi-modal dimension of ordinary aesthetic experience

Highlights

- Are individual differences in aesthetic impressions systematic or arbitrary?

- We measured "taste typicality" by comparing individuals' tastes to the average taste

- We found that visual and auditory taste typicality was systematically correlated

- Taste typicality was also the primary way people's tastes differed from each other

Authors

Yi-Chia Chen, Andrew Chang, Monica D. Rosenberg, Derek Feng, Brian J. Scholl, Laurel J. Trainor

Correspondence

yichiachen@g.ucla.edu

In brief

Chen et al. show that people's aesthetic tastes are not arbitrarily different from each other in different sensory modalities but vary primarily along only a single dimension across sights and sounds: how similar a person's taste is to the average taste. People who have atypical taste for images also tend to have atypical taste for sounds.

Report

"Taste typicality" is a foundational and multi-modal dimension of ordinary aesthetic experience

Yi-Chia Chen,[1,2,10,*] Andrew Chang,[3,4] Monica D. Rosenberg,[5] Derek Feng,[6] Brian J. Scholl,[7] and Laurel J. Trainor[3,8,9]
[1]Department of Psychology, Harvard University, Cambridge, MA 02138, USA
[2]Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA
[3]Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON L8S 4K1, Canada
[4]Department of Psychology, New York University, New York, NY 10003, USA
[5]Department of Psychology, University of Chicago, Chicago, IL 60637, USA
[6]Department of Statistics and Data Science, Yale University, New Haven, CT 06511, USA
[7]Department of Psychology, Yale University, New Haven, CT 06511, USA
[8]McMaster Institute for Music and the Mind, McMaster University, Hamilton, ON L8S 4K1, Canada
[9]Rotman Research Institute, Baycrest Hospital, Toronto, ON M6A 2E1, Canada
[10]Lead contact
*Correspondence: yichiachen@g.ucla.edu
https://doi.org/10.1016/j.cub.2022.02.039

SUMMARY

Aesthetic experience seems both regular and idiosyncratic. On one hand, there are powerful regularities in what we tend to find attractive versus unattractive (e.g., beaches versus mud puddles).[1–4] On the other hand, our tastes also vary dramatically from person to person:[5–8] what one of us finds beautiful, another might find distasteful. What is the nature of such differences? They may in part be arbitrary—e.g., reflecting specific past judgments (such as liking red towels over blue ones because they were once cheaper). However, they may also in part be systematic—reflecting deeper differences in perception and/or cognition. We assessed the systematicity of aesthetic taste by exploring its *typicality* for the first time across seeing and hearing. Observers rated the aesthetic appeal of ordinary scenes and objects (e.g., beaches, buildings, and books) and environmental sounds (e.g., doorbells, dripping, and dialtones). We then measured "taste typicality" (separately for each modality) in terms of the similarity between each individual's aesthetic preferences and the population's average. The data revealed two primary patterns. First, taste typicality was not arbitrary but rather was correlated to a moderate degree across seeing and hearing: people who have typical taste for images also tend to have typical taste for sounds. Second, taste typicality captured most of the explainable variance in people's impressions, showing that it is the primary dimension along which aesthetic tastes systematically vary.

RESULTS

One of the most prominent aspects of aesthetic experience is the degree to which people's aesthetic tastes differ. People regularly and consistently disagree about what looks more or less appealing, and these differences clearly transcend mere instances of uncertainty or ambiguity: often, our own tastes seem so clear and direct that it can be perplexing (if not maddening) to learn that others do not share them. Are these differences truly arbitrary?

In the current study, we explore the nature of individual aesthetic tastes by asking how different domains of aesthetic taste (e.g., involving different sensory modalities) relate to each other. Are they independent, or do we have a general aesthetic orientation that affects them all? We determined and quantified what we call people's "taste typicality" across multiple sensory modalities—focusing in particular on seeing and hearing. And while doing so, we sought to draw contrasts with past work in four ways.

First, past studies have compared individuals' tastes to some sort of expert standard.[9–12] Here, we correlate individuals' preferences to the *average* preferences of the population they come from—seeking to determine not how "correct" or sophisticated those preferences are but simply how typical they are. Where past studies have computed such correlations[6,8] to assess overall consensus, here we use them to measure individual aesthetic tastes. Second, past studies of individual aesthetic preferences have often used specialized stimuli, such as artwork,[13–15] music,[16] architecture,[13] cars,[17] or abstract visual shapes.[18–20] Here, we measured aesthetic preferences for maximally general types of "ordinary" stimuli—both visual scenes (such as a brick wall, some clouds, and an office) and environmental sounds (such as a hair dryer, typing, and thunder), as shown in Figure 1. As a result, our experiments do not attempt to address the heights of aesthetic beauty,[21–23] but they instead capture the type of aesthetic impressions that permeate our everyday lives. Such "ordinary" stimuli have been explored in empirical aesthetics more generally,[22,24]

A

B

Cellphone Vibrating	Rubbing Hands	Boiling Water	Typing
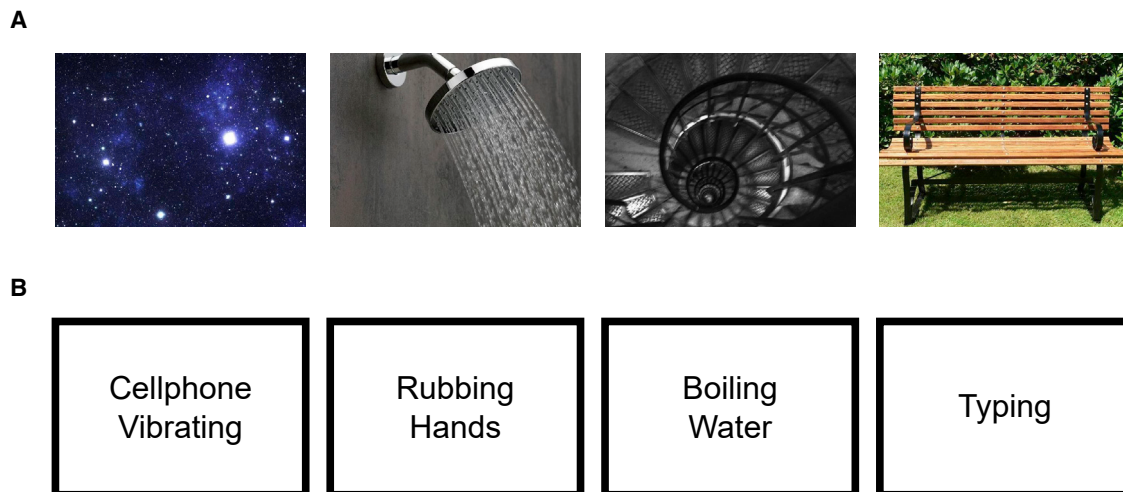
Figure 1. Example stimuli
(A) Example visual stimuli.
(B) Descriptions of example auditory stimuli.

but not in studies of aesthetic taste—perhaps because of strong associations between taste and sophistication. One may be accused of having "bad taste in art," but not "bad taste in clouds"! Third and relatedly, the focus on everyday aesthetics not only guided the kind of stimuli used here but also decided the kind of aesthetic experiences measured. While some researchers have treated aesthetics as more refined and complex than "mere" preferences,[25–27] it seems clear that the core of aesthetic experiences are the preferences themselves, at least in our everyday lives. Thus, our study does not seek to contribute to theoretical discussions of how to best define art and aesthetics but instead simply focuses on the aesthetic experiences that arise when we encounter something that we find appealing. Finally, many past studies have focused on the specific perceptual,[5,28–30] cognitive,[5,31,32] and/or cultural factors[33] that contribute to aesthetic experiences. Here, we took a different strategy to explore the holistic patterns of individuals' aesthetic tastes.

We measured taste typicality in a straightforward way: by correlating each individual's preference for each item with the population mean. We aimed to address two questions. First, how does typicality vary (or not) across independent stimulus modalities: if you have highly typical/atypical preferences for visual stimuli, will you also have highly typical/atypical preferences for environmental sounds? Second, how fundamental is taste typicality in the first place: how much of the overall variance in preference ratings can be captured by this factor?

Two hundred observers (100 in a primary experiment and 100 in a direct replication) participated in an online study with their own web browser and headphones (tested with a screening procedure that preceded the main experiment) and rated how visually or aurally appealing 88 images and 88 sounds (Figures 1 and S1) were on a scale from 1 ("very not appealing") to 6 ("very appealing"). Test-retest reliabilities were assessed by repeating a subset of the stimuli at the end of each experiment. A demonstration of the experiments is accessible online at https://yi-chia-chen.github.io/taste-typicality-demo-expt/.

Taste typicality for each modality

Each observer's initial ratings for all images were first converted to (within-subject) z-scores, and each observer's visual taste typicality score was then calculated as the correlation between that observer's Z scores and the averaged Z scores across all other observers (not including the observer in question) for each image. This same procedure was then used with the ratings of environmental sounds to produce an auditory taste typicality score for each observer. These two modality-specific taste typicality scores were then analyzed both for a primary experiment and a direct replication.

We first confirmed that these scores were generalizable across stimuli by performing a split-half reliability analysis. Within each modality, each observer's stimuli were randomly split into two halves, separate taste typicality scores were computed for each half, and then these two scores were correlated with each other (corrected with a Spearman-Brown formula). The average reliability across 1,000 random splits for each observer was always above 0.5, revealing that the taste typicality scores were not fully dependent on particular stimuli (primary images: $M = 0.678$, SD = 0.050; primary sounds: $M = 0.636$, SD = 0.054; replication images: $M = 0.585$, SD = 0.063; replication sounds: $M = 0.695$, SD = 0.047).

The distributions of taste typicality scores for each modality are depicted in Figures 2A and 2B as a smoothed histogram (with kernel density estimation). Inspection of these distributions reveals (1) that many observers had highly typical aesthetic tastes (as indicated by the rightmost parts of each distribution), (2) that many other observers' aesthetic tastes were highly idiosyncratic (as indicated by the leftmost parts of each distribution), but (3) that no observer in this dataset had aesthetic tastes that were anticorrelated with the average taste (as indicated by the absence of any negative typicality scores). In general, most observers had typicality scores for each modality that ranged from 0.5 to 0.8 (primary images: $M = 0.664$ with range [0.315, 0.839], SD = 0.107; primary sounds: $M = 0.570$ [0.135, 0.815], SD = 0.127; replication

A   Primary Experiment

B   Replication
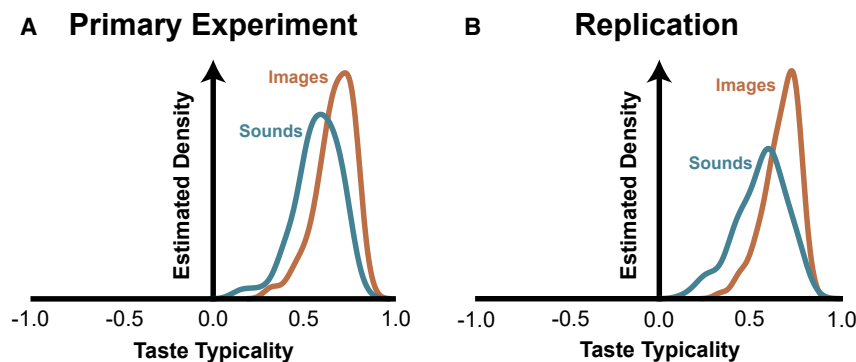
Figure 2. Distributions of taste typicality scores

The distributions of taste typicality scores were plotted with kernel density estimation using Gaussian distributions and with bandwidth selected based on Scott's rule. The distributions from (A) the primary experiment and (B) the replication were plotted for images (orange) and sounds (cyan) separately.

images: $M = 0.663$ [0.337, 0.817], SD = 0.096; replication sounds: $M = 0.555$ [0.160, 0.813], SD = 0.137).

Taste typicality across modalities

The relationship between taste typicality in seeing and hearing was then assessed simply by correlating the two scores across all observers. If people have independent aesthetic *tastes* in different modalities, then we would expect no relationship between the two scores. But if there is also a more general aesthetic taste typicality that is shared across modalities, then we would expect some relationship between the two scores. The results of this correlation, as depicted by the scatterplots in Figures 3A and 3B, clearly support this latter possibility: higher visual taste typicality scores tended to co-occur with higher auditory taste typicality scores (primary: $r_{raw} = 0.241$, $CI = [0.047, 0.418]$, $r_{corrected} = 0.367$, p = 0.016; replication: $r_{raw} = 0.330$, $CI = [0.143, 0.494]$, $r_{corrected} = 0.518$, p < 0.001, with Spearman's correction for attenuation using the split-half reliabilities; all tests reported in this study are two-tailed tests). This reliable positive correlation also survived after controlling for the observers' test-retest reliabilities; this ensures that shared atypicality across modalities cannot simply be reduced to unreliable patterns of responses, which, for example, would be the case if some observers (but not others) were simply responding randomly (primary: $r_{partial} = 0.209$ [0.013, 0.389], p = 0.043; replication: $r_{partial} = 0.302$ [0.112, 0.471], p = 0.003). In other words, people with typical preferences for visual stimuli do also tend to have typical preferences for environmental sounds.

### How central is typicality to aesthetic taste?

Having identified an attribute of taste typicality that is shared to some degree across modalities, we can further ask how important this factor is to individual aesthetic taste. Taste typicality could be a central component of aesthetic taste, such that it could account for much of the systematic variance in aesthetic taste more broadly. But it could also be just one of many general factors that each capture important parts of the underlying variance in aesthetic taste (e.g., preferences for complexity versus simplicity or regularity versus chaos). To find out, we ran principal components analyses on the ratings (after *Z* scoring within observers and modalities) over observers to extract those dimensions that predict variation across observers (see STAR Methods for details). The results of these analyses are plotted in Figures 4A and 4B for images and sounds separately and for all stimuli regardless of modality.

The central result of this analysis is that a single factor dominates aesthetic taste. This is apparent in the figures in the stark L-shaped patterns. In contrast, if there were many orthogonal factors that collectively comprised taste, then these lines would each have had a more gradual decrease as a function of the principal component number. Indeed, the first principal component always accounted for at least one third of the overall variance (primary images: 46.3%; primary sounds: 35.3%; primary all: 40.3%; replication images: 45.9%; replication sounds: 33.9%; replication all: 39.6%), and the second principal component never accounted for more than 7.5% additional variance (primary images: 4.9%; primary sounds: 5.3%; primary all: 3.6%; replication images: 5.0%; replication sounds: 7.4%; replication all: 4.2%).

Critically, the first principal component appeared to be precisely taste typicality, given that these two dimensions were exceptionally highly correlated (primary images: $r > 0.999$; primary sounds: $r > 0.999$; primary all: $r_v = 0.781$, $r_a = 0.794$; replication images: $r > 0.999$; replication sounds: $r = 0.999$; replication all: $r_v = 0.787$, $r_a = 0.775$; all p < 0.001). Thus, taste typicality is not just one way (among many) that people differ in their aesthetic tastes, but it is rather the primary way in which our tastes differ.

DISCUSSION

The stimuli used in many studies of aesthetic appreciation are *beautiful*—sometimes encompassing ravishing works of art[34,35] or arresting musical passages.[36,37] This is not so for the current study, which instead employed stimuli one might encounter during everyday life. (We suspect that few would find the bench from Figure 1 to be ravishing.) Nevertheless, observers often agreed with each other about how appealing these images and sounds were, and we suspect that this is a hallmark of "everyday aesthetics"—the sort of aesthetic experiences one has, not when listening to a concert, but when walking back to your car afterward. The use of such ordinary stimuli also rules out other potential concerns. When studying taste typicality, in particular, some people may just want to be—or seem—*different* and so intentionally respond in ways that set them apart from the crowd. Here, in contrast, not a single observer's aesthetic preferences were anticorrelated with the population mean. This may be because of the nature of our stimuli: you can't "fight the crowd" if you have no idea what the crowd would think in the first place, and we don't generally have stereotypes about, for example, the degree to which the noise of two rubbing hands is appealing. In the present study, we used aesthetic judgments of such stimuli in order to address two particular questions.
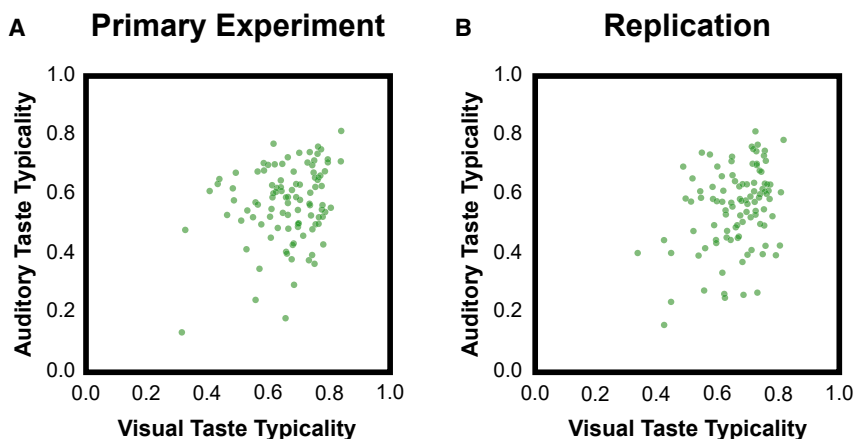
A   Primary Experiment

B   Replication

We first asked whether one's taste typicality for seeing was fully independent of one's taste typicality for hearing. The answer was clearly no: these two dimensions were robustly correlated with each other to a moderate degree—such that it is possible to predict taste typicality in one modality from taste typicality in another modality. We next asked how central taste typicality is to individual differences in aesthetic taste: even when robust, it might nevertheless be a peripheral factor in aesthetic taste, which is eclipsed by other factors that are able to explain more variance in aesthetic tastes. Here, the data provided a clear and powerful answer: taste typicality appears to be the primary determinant of individual differences in aesthetic preferences—in both seeing and hearing—with no other factor able to explain even one fourth as much variance.

These results might initially seem to be at odds with past research, which has identified many specific factors that appear to drive aesthetic impressions. For example, past work has shown that people prefer stimuli that are blue,[2] curvy,[24] symmetrical,[18] inward-facing,[38,39] moderately complex,[40] balanced,[41] typical[42] (but see Vogel, Ingendahl, and Winkielman[43]), and presented in a canonical size.[44–46] One possibility is that such properties are orthogonal to taste typicality, and they instead contribute to other principal components of the variance in aesthetic impressions—although, this would mean that they could collectively explain no more than 7.5% of such variance (Figure 4). A second possibility, however, is that these other factors are central after all, insofar as they all contribute to the dimension of typicality. In this case, however, our data add to previous work by demonstrating indirectly that the individual tastes for these seemingly independent properties (e.g., curviness and complexity) must in fact be related—since for them to collectively constitute taste typicality, they would have to be strongly correlated with each other (since otherwise, they would be split into multiple principal components).[7] This same lesson also applies to other less perceptual factors. Others have attempted to explain individual differences in aesthetic appreciation by appealing to factors such as personality[47] (but see McManus, Cook, and Hunt[48]), expertise,[7,49] life history,[2,50] and neuroanatomy.[16,37] For example, more art education leads to less typical preferences for harmony and symmetry[7,51,52] (potentially through social learning[53]). However, at least for the "ordinary" stimuli explored in here, these other seemingly disparate factors could only play a substantive role if they were also correlated

with each other—such that they too contribute to the first principal component of the variability in aesthetic impressions.

Of course, it will come as a surprise to nobody that people vary in how typical their aesthetic tastes are. Most of us know others whose tastes in music or film are either "mainstream" or "alternative." The present work demonstrates that this form of typicality is fundamental to our aesthetic impressions in two ways: it is a broad factor that operates to some degree across sensory modalities, and it explains more variability in aesthetic impressions than any other single factor does.

These discoveries open the door for a new way to explore aesthetic experiences. For example, in addition to asking what factors give rise to typical aesthetic experiences, we may ask how people come to share many aesthetic responses and how an individual comes to deviate from that. While the former can arise from interactions between the stimulus properties and the evolutionary goals[54] or common experiences[55] of humans, our findings demand a different kind of answer to the latter question. Given the idiosyncratic ways in which people's tastes differ from what is typical—and the similar degree of such deviations across different modalities within each individual—the responsible mechanisms must operate in a broad and coherent way over multiple domains rather than being reducible to any processes that lead to specific instances of preferences, such as familiarity with specific stimuli. We hope that this work will spur future research on just what constitutes these mechanisms and how they give rise to systematic individual aesthetic tastes.

Conclusion

This study explored how people's aesthetic tastes for ordinary visual stimuli relate to those for ordinary environmental sounds. Using a measure of taste typicality to compare individuals to the population they came from, we made two main discoveries. First, we found that people's taste typicalities were robustly correlated between seeing and hearing—such that the more typical one's taste was for visual stimuli, the more typical one's taste was for environmental sounds. Second, taste typicality was the primary way people's aesthetic preferences differed from each other, with no other factor explaining even one fourth as much variance.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - ○ Lead contact

A   **Primary Experiment**

B   **Replication**
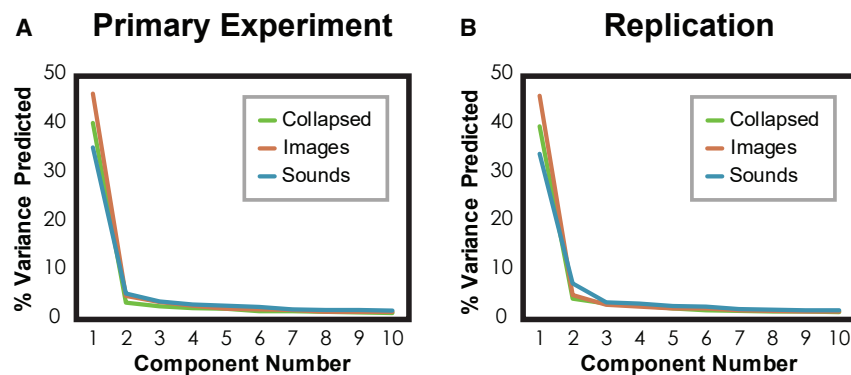
Figure 4. Results from principal component analyses

The percentage of variance explained by the first 20 components extracted from ratings for images (orange), for sounds (cyan), and for all stimuli (green) after *Z* transformation in (A) the primary experiment and (B) the replication.

○ Materials availability
○ Subject details
○ Data and code availability
● METHOD DETAILS
  ○ General procedure
  ○ Volume adjustment procedure
  ○ Headphone screening procedure
  ○ Rating procedure
  ○ Debriefing question procedure
  ○ Visual stimuli
  ○ Auditory stimuli
● QUANTIFICATION AND STATISTICAL ANALYSIS
  ○ Observer exclusions
  ○ General analysis
  ○ Additional analysis
  ○ Principal component analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.02.039.

AUTHOR CONTRIBUTIONS

All authors designed the research and wrote the manuscript. Y.-C.C., A.C., and M.D.R. prepared the materials. Y.-C.C. and A.C. conducted the experiments. Y.C., A.C., and D.F. analyzed the data.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. McManus, I.C. (1980). The aesthetics of simple figures. Br. J. Psychol. *71*, 505–524.

2. Palmer, S.E., and Schloss, K.B. (2010). An ecological valence theory of human color preference. Proc. Natl. Acad. Sci. USA *107*, 8877–8882.

3. Sherman, A., Grabowecky, M., and Suzuki, S. (2015). In the working memory of the beholder: Art appreciation is enhanced when visual complexity is compatible with working memory. J. Exp. Psychol. Hum. Percept. Perform. *41*, 898–903.

4. Van Geert, E., and Wagemans, J. (2021). Order, complexity, and aesthetic preferences for neatly organized compositions. Psychology of Aesthetics, Creativity, and the Arts *15*, 484–504.

5. Iigaya, K., Yi, S., Wahle, I.A., Tanwisuth, K., and O'Doherty, J.P. (2021). Aesthetic preference for art can be predicted from a mixture of low- and high-level visual features. Nat. Hum. Behav. *5*, 743–755.

6. Leder, H., Goller, J., Rigotti, T., and Forster, M. (2016). Private and shared taste in art and face appreciation. Front. Hum. Neurosci. *10*, 155.

7. Palmer, S.E., and Griscom, W.S. (2013). Accounting for taste: individual differences in preference for harmony. Psychon. Bull. Rev. *20*, 453–461.

8. Vessel, E.A., Maurer, N., Denker, A.H., and Starr, G.G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. Cognition *179*, 121–131.

9. Wallisch, P., and Whritner, J.A. (2017). Strikingly low agreement in the appraisal of motion pictures. Projections *11*, 102–120.

10. Götz, K.O., Lynn, R., Borisy, A.R., and Eysenck, H.J. (1979). A new visual aesthetic sensitivity test: I. Construction and psychometric properties. Percept. Mot. Skills *49*, 795–802.

11. Myszkowski, N., and Storme, M. (2017). Measuring "good taste" with the visual aesthetic sensitivity test-revised (VAST-R). Pers. Individ. Dif. *117*, 91–100.

12. Corradi, G., Chuquichambi, E.G., Barrada, J.R., Clemente, A., and Nadal, M. (2020). A new conception of visual aesthetic sensitivity. Br. J. Psychol. *111*, 630–658.

13. Vessel, E.A., Isik, A.I., Belfi, A.M., Stahl, J.L., and Starr, G.G. (2019). The default-mode network represents aesthetic appeal that generalizes across visual domains. Proc. Natl. Acad. Sci. USA *116*, 19155–19164.

14. Eysenck, H.J. (1940). The general factor in aesthetic judgements. Br. J. Psychol. *31*, 94–102.

15. Opacic, T., Stevens, C., and Tillmann, B. (2009). Unspoken knowledge: implicit learning of structured human dance movement. J. Exp. Psychol. Learn. Mem. Cogn. *35*, 1570–1577.

16. Belfi, A.M., and Loui, P. (2020). Musical anhedonia and rewards of music listening: current advances and a proposed model. Ann. N Y Acad. Sci. *1464*, 99–114.

17. Carbon, C.C. (2010). The cycle of preference: long-term dynamics of aesthetic appreciation. Acta Psychol. (Amst.) *134*, 233–244.

18. Jacobsen, T., and Höfel, L. (2002). Aesthetic judgments of novel graphic patterns: analyses of individual judgments. Percept. Mot. Skills 95, 755–766.

19. Jacobsen, T. (2004). Individual and group modelling of aesthetic judgment strategies. Br. J. Psychol. 95, 41–56.

20. McManus, I.C. (1980). The aesthetics of simple figures. Br. J. Psychol. 71, 505–524.

21. Brielmann, A.A., and Pelli, D.G. (2017). Beauty requires thought. Curr. Biol. 27, 1506–1513.e3.

22. Brielmann, A.A., and Pelli, D.G. (2019). Intense beauty requires intense pleasure. Front. Psychol. 10, 2420.

23. Makin, A.D.J. (2017). The gap between aesthetic science and aesthetic experience. J. Conscious. Stud. 24, 184–213.

24. Bar, M., and Neta, M. (2006). Humans prefer curved visual objects. Psychol. Sci. 17, 645–648.

25. Wassiliwizky, E., and Menninghaus, W. (2021). Why and how should cognitive science care about aesthetics? Trends Cogn. Sci. 25, 437–449.

26. Palmer, S.E., Schloss, K.B., and Sammartino, J. (2013). Visual aesthetics and human preference. Annu. Rev. Psychol. 64, 77–107.

27. Pearce, M.T., Zaidel, D.W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., and Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. Perspect. Psychol. Sci. 11, 265–279.

28. Isherwood, Z.J., Clifford, C.W.G., Schira, M.M., Roberts, M.M., and Spehar, B. (2021). Nice and slow: Measuring sensitivity and visual preference toward naturalistic stimuli varying in their amplitude spectra in space and time. Vision Res. 181, 47–60.

29. Chen, Y.-C., and Scholl, B.J. (2014). Seeing and liking: biased perception of ambiguous figures consistent with the "inward bias" in aesthetic preferences. Psychon. Bull. Rev. 21, 1444–1451.

30. Forman, I.R., Chen, Y.-C., Scholl, B.J., and Alvarez, G.A. (2021). The center cannot hold: Variations of frame width help to explain the "inward bias" in aesthetic preferences. Atten. Percept. Psychophys. 83, 2151–2158.

31. Van de Cruys, S. (2017). Affective value in the predictive mind. In Philosophy and Predictive Processing, T. Metzinger, and W. Wiese, eds. (MIND Group), pp. 24:1–2421.

32. Carbon, C.C. (2011). Cognitive mechanisms for explaining dynamics of aesthetic appreciation. Iperception 2, 708–719.

33. Masuda, T., Gonzalez, R., Kwan, L., and Nisbett, R.E. (2008). Culture and aesthetic preference: comparing the attention to context of East Asians and Americans. Pers. Soc. Psychol. Bull. 34, 1260–1275.

34. Vissers, N., Moors, P., Genin, D., and Wagemans, J. (2020). Exploring the role of complexity, content and individual differences in aesthetic reactions to semi-abstract art photographs. Art & Perception 8, 89–119.

35. Vessel, E.A., Starr, G.G., and Rubin, N. (2012). The brain on art: intense aesthetic experience activates the default mode network. Front. Hum. Neurosci. 6, 66.

36. Cheung, V.K.M., Harrison, P.M.C., Meyer, L., Pearce, M.T., Haynes, J.D., and Koelsch, S. (2019). Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity. Curr. Biol. 29, 4084–4092.e4.

37. Sachs, M.E., Ellis, R.J., Schlaug, G., and Loui, P. (2016). Brain connectivity reflects human aesthetic responses to music. Soc. Cogn. Affect. Neurosci. 11, 884–891.

38. Palmer, S.E., Gardner, J.S., and Wickens, T.D. (2008). Aesthetic issues in spatial composition: effects of position and direction on framing single objects. Spat. Vis. 21, 421–449.

39. Chen, Y.-C., Colombatto, C., and Scholl, B.J. (2018). Looking into the future: An inward bias in aesthetic experience driven only by gaze cues. Cognition 176, 209–214.

40. Martindale, C., Moore, K., and West, A. (1988). Relationship of preference judgments to typicality, novelty, and mere exposure. Empir. Stud. Arts 6, 79–96.

41. Locher, P., Cornelis, E., Wagemans, J., and Stappers, P.J. (2001). Artists' use of compositional balance for creating visual displays. Empir. Stud. Arts 19, 213–227.

42. Martindale, C., and Moore, K. (1988). Priming, prototypicality, and preference. J. Exp. Psychol. Hum. Percept. Perform. 14, 661–670.

43. Vogel, T., Ingendahl, M., and Winkielman, P. (2021). The architecture of prototype preferences: Typicality, fluency, and valence. J. Exp. Psychol. Gen. 150, 187–194.

44. Konkle, T., and Oliva, A. (2011). Canonical visual size for real-world objects. J. Exp. Psychol. Hum. Percept. Perform. 37, 23–37.

45. Linsen, S., Leyssen, M.H., Sammartino, J., and Palmer, S.E. (2011). Aesthetic preferences in the size of images of real-world objects. Perception 40, 291–298.

46. Chen, Y.-C., Deza, A., and Konkle, T. (2020). How big should this object be? Perceptual influences on viewing-size preferences. Journal of Vision 20, 428.

47. Chamorro-Premuzic, T., Reimers, S., Hsu, A., and Ahmetoglu, G. (2009). Who art thou? Personality predictors of artistic preferences in a large UK sample: the importance of openness. Br. J. Psychol. 100, 501–516.

48. McManus, I.C., Cook, R., and Hunt, A. (2010). Beyond the golden section and normative aesthetics: Why do individuals differ so much in their aesthetic preferences for rectangles? Psychology of Aesthetics, Creativity, and the Arts 4, 113–126.

49. Silvia, P.J., and Barona, C.M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. Empir. Stud. Arts 27, 25–42.

50. Van de Cruys, S., and Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. Iperception 2, 1035–1062.

51. Leder, H., Tinio, P.P.L., Brieber, D., Kröner, T., Jacobsen, T., and Rosenberg, R. (2019). Symmetry is not a universal law of beauty. Empir. Stud. Arts 37, 104–114.

52. Weichselbaum, H., Leder, H., and Ansorge, U. (2018). Implicit and explicit evaluation of visual symmetry as a function of art expertise. Iperception 9, 2041669518761464.

53. Boddez, Y., Descheemaeker, M., Mertens, G., Truyts, A., and Van de Cruys, S. (2019). Like what you see: Generalization of social learning determines art appreciation. Acta Psychol. (Amst.) 196, 18–25.

54. Orians, G.H., and Heerwagen, J.H. (1992). Evolved responses to landscapes. In The Adapted Mind: Evolutionary Psychology and the Generation of Culture, J.H. Barkow, L. Cosmides, and J. Tooby, eds. (Oxford University Press), pp. 555–579.

55. Moon, C., Cooper, R.P., and Fifer, W.P. (1993). Two-day-olds prefer their native language. Infant Behav. Dev. 16, 495–500.

56. Norman-Haignere, S., Kanwisher, N.G., and McDermott, J.H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. Neuron 88, 1281–1296.

57. Crump, M.J., McDonnell, J.V., and Gureckis, T.M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. PLoS ONE 8, e57410.

58. Woods, K.J.P., Siegel, M.H., Traer, J., and McDermott, J.H. (2017). Headphone screening to facilitate web-based auditory experiments. Atten. Percept. Psychophys. 79, 2064–2072.

59. Chandler, J., Mueller, P., and Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. Behav. Res. Methods 46, 112–130.

60. Kittur, A., Chi, E.H., and Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. In Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, M. Czerwinski, and A. Lund, eds. (ACM), pp. 453–456.

61. Zhu, D., and Carterette, B. (2010). An analysis of assessor behavior in crowdsourced preference judgments. Proceedings of the ACM SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation, 21–26.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed data	This paper	https://doi.org/10.17605/OSF.IO/T925Q
Auditory stimuli	PacDV free sound effects	https://www.pacdv.com/sounds/
Auditory stimuli	Norman-Haignere et al., 2015[56]	https://doi.org/10.1016/j.neuron.2015.11.035
Software and algorithms		
Experiment code	This paper	https://doi.org/10.17605/OSF.IO/T925Q
Analysis code	This paper	https://doi.org/10.17605/OSF.IO/T925Q
Other		
Full set of visual and auditory stimuli	This paper	https://doi.org/10.17605/OSF.IO/T925Q
Demonstration experiment	This paper	https://yi-chia-chen.github.io/taste-typicality-demo-expt/

RESOURCE AVAILABILITY

Lead contact
Further information and requests should be directed to and will be fulfilled by the lead contact, Yi-Chia Chen (yichiachen@g.ucla.edu).

Materials availability
This study did not generate new unique material.

Subject details
For each experiment, 100 human observers (gender and age not recorded) participated in a 25-min experiment in exchange for monetary compensation or course credit. The sample size was predetermined based on informal pilot experiments, and the replication confirmed that this sample size provided sufficient power to detect relevant effects. The primary experiment was run through Amazon Mechanical-Turk (MTurk).[57] All observers were in the U.S. or Canada, had an MTurk task approval rate of at least 85%, and had previously completed at least 50 MTurk tasks. We also prevented observers who had seen the same set of stimuli in other studies from participating. The replication was run online with students at McMaster University, Harvard University, and Yale University. All observers reported normal hearing and normal or corrected-to-normal visual acuity, normal color vision, and no history of a vision disorder. The study was approved by McMaster Research Ethics Board, Harvard University Area Institutional Review Board, and Yale University Institutional Review Board.

Data and code availability

- All stimuli and raw data have been deposited at an OSF repository and are publicly available as of the date of publication. Access link is listed in the key resources table.
- All original experiment and analysis code have been deposited at the same OSF repository and are publicly available as of the date of publication. Access link is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

General procedure
After agreeing to participate, observers were redirected to a website where stimulus presentation and data collection were controlled via custom software written in HTML, CSS, JavaScript, and PHP. Since the experiment was rendered on observers' own web browsers, viewing distance, screen size, and display resolutions could vary dramatically, and so we report visual stimulus dimensions below using pixel (px) values. To avoid compatibility issues, we asked observers to use any browsers other than Internet Explorer, and blocked participation with phones or tablets. Also, to make sure sounds were presented with acceptable quality, we asked observers to use headphones (instead of loudspeakers) during the experiments. This requirement was directly tested in the experiment. Observers who didn't follow the instructions and participated without headphones were excluded and replaced. Because it was not feasible to measure the actual volume on each observer's headphones, we only report relative loudness.

The experiments had 4 parts: Volume adjustment, headphone screening, ratings, and debriefing questions. The first two parts ensured that the sounds were presented properly in the third part for rating. A demonstration of the experiments can be viewed online here: https://yi-chia-chen.github.io/taste-typicality-demo-expt/.

Volume adjustment procedure
The observers were asked to put on headphones and make sure that they could hear sounds. They were then asked to set the computer volume to about 20% of maximum and make sure that the audio device was not muted. Next, a calibration sound (a 2 s-long noise with loudness and spectrum matching the average of all sounds used later for rating) was played and the observers adjusted the volume so that the sound was at a loud but comfortable level. They could play the sound as many times as needed by clicking on a button.

Headphone screening procedure
Observers performed a simple intensity discrimination task (3-alternative forced choice) to detect their headphone use.[58] They hit a button to hear three tones separated by silences (played only once) and judged which sound was the quietest. The three 1 s 200-Hz pure tones (with 100 ms on- and off-ramps; including a binaurally in-phase loud tone, an antiphase loud tone, and an in-phase quiet tone of -6 dB) were presented in a random order. In the anti-phase loud tone, the two channels (left and right) emitted a high-pressure peak and a low-pressure trough at the same time, leading to attenuation of sound by phase-cancellation if played through loudspeakers. This cancellation does not occur if the sounds are heard through headphones (since the sound waves from the two channels do not encounter each other in the open air). Thus, the task can only be accurately performed if the observers are wearing headphones. Observers who performed worse than 6 out of 7 trials correct (after 1 practice trial) were prevented from continuing. This criterion leads to minimum miss rates and low false alarm rates at detecting non-headphone users.[58]

Rating procedure
Observers were asked to rate from 1-6 how "visually appealing" each image was in a block, and how "aurally appealing" each sound was in another block (with the block order randomized, and the two blocks separated by a self-paced break). The scale was anchored with 6 as "Very appealing" and 1 as "Very not appealing." Both images and sounds were presented for 2 s only once, and the observers were given unlimited time to respond. (Observers could respond during image presentation but were only allowed to respond after the sound completed playing.) Before the rating trials started, observers were tested for their understanding of the instructions with a multiple-choice question, and the instructions repeated if they answered incorrectly. Each block started with two practice trials, for which the ratings were not recorded. 88 stimuli followed in random order. Afterward, 20 random stimuli were repeated in the primary experiment and 35 in the replication in each block for measuring test-retest reliability of individuals' ratings. These repeating trials were not included in the main analyses.

Debriefing question procedure
After the rating phase, observers rated how often they spontaneously noticed how visually and aurally appealing (or unappealing) things are in day-to-day life (on a scale from 1 to 7, where 1 means "very seldom" and 7 means "very often"). They were also asked to report any problems or if they found the procedure unclear. In the replication, they were additionally asked if they had completed the experiment seriously throughout (without randomly clicking through any part of the experiment).

Visual stimuli
Images were collected through Google Image search, with keywords generated with online random word generators. We looked at the top five results with the "Size" search setting set to "Large." Only images passing the following criteria were included: Related to the keyword, easily interpretable as a real photograph without visible alterations, larger than 750 × 550 px, has at least one distinct object (excluding uniform textures), with content that is not obviously emotional, and does not include any (realistic or cartoony) people, body parts, animals, symbols (e.g., a brand mark or dollar sign), or text. We chose to exclude uniform textures, emotional and animate objects, and communicative symbols because they likely engage specialized processes. The images were then resized to their respective smallest size that were still larger than 750 × 550 px and cropped to retain a random 650 × 450 px region. This random cropping is used to diversify the framing in our image set, since photographs online were often selectively framed by people in ways that may involve certain biases. Informal pilot studies were run to select 88 from 200 images that spanned the full range of aesthetic values, along with 2 additional neutral images for practice trials (see Figure S1 for the final selection).

Auditory stimuli
We selected the 88 environmental sounds from a database[56] and the 2 sounds used for practice trials from a sound-effects website (https://www.pacdv.com/sounds/). Only sounds passing the following criteria were used: not interpretable as music with melody, with content that is not obviously emotional, and does not include any human voice, speech, or animal call. These exclusions were based on similar considerations about engaging specialized processes. Each sound was 2 s long, with perceptual loudness equalized using the replayGain 1.0 algorithm (implemented with Audacity 2.1.1).

All stimuli (labeled with their respective keywords or names) along with all the raw data have been deposited at the public OSF repository listed in the key resources table.

QUANTIFICATION AND STATISTICAL ANALYSIS

Observer exclusions

56 observers in the primary experiment were excluded and replaced (based on criteria decided before data collection began,[59] that were used in both the primary experiment and the replication; with some observers triggering more than 1 criterion): 2 observers who reported technical problems, 2 observers who had browser viewports smaller than 650 × 550 px, 6 observers who answered the question about the instructions incorrectly more than once, 12 observers who had more than 4 abnormal RTs in either block (RTs shorter than 300 ms in the visual block, or longer than 2 min in both blocks), 5 observers who gave uniform or temporally regular ratings (which showed that they were obviously not following instructions), 23 observers with lower than 0.5 test-retest reliability with images, and 44 observers with lower than 0.5 test-retest reliability with sounds. We also excluded and replaced 1 observer who reported having high-functioning autism spectrum disorder (which was not a predetermined criterion). The predetermined exclusion criteria were in place to ensure that the data quality was comparable to in-lab studies.[60,61]

Forty-three observers in the replication were excluded and replaced according to the same criteria: 2 observers who had small browser viewports, 4 observers who failed the instructions question more than once, 6 observers who had uniform or temporally regular ratings, and 19 observers who had low test-retest reliability with sounds. At the same time, 6 observers were excluded and replaced based on reporting not being serious in the additional debriefing question.

These exclusion criteria led to reasonable test-retest reliabilities in all conditions: Primary images: $M = 0.856$, $range = [0.565, 0.994]$, SD = 0.105; Primary sounds: $M = 0.787$, $range = [0.515, 1.000]$, SD = 0.114; Replication images: $M = 0.857$, $range = [0.554, 0.984]$, SD = 0.086; Replication sounds: $M = 0.756$, $range = [0.501, 0.938]$, SD = 0.105.

General analysis

All analyses reported here were conducted with customized Python code using the pandas library, which have been deposited at the public OSF repository listed in the key resources table. Statistical test significance was assessed with $\alpha = 0.05$. The specific tests used are reported in the Results section.

Additional analysis

Our primary analyses were conducted using z-scores since different observers may use the scale in different ways (such that the very same impression might receive a "4" rating from one observer, but a "6" from another). But in fact, the results of all analyses reported in the Results section remained qualitatively identical when using the initial absolute ratings directly. And these absolute ratings spanned the full scale for nearly every observer (195/200 observers), for the Primary images ($M = 3.6$, SD = 1.7, $Range = [1, 6]$), Primary sounds ($M = 2.7$, SD = 1.5, $Range = [1, 6]$), Replication images ($M = 3.5$, SD = 1.8, $Range = [1, 6]$), and Replication sounds ($M = 3.0$, SD = 1.6, $Range = [1, 6]$). The qualitative patterns of results also remained the same when we excluded the 5 observers who did not use the full 6-point scale.

Principal component analysis

We conducted the principal components analyses (PCA) to reduce dimensionality in the multi-dimensional space where each axis represents the ratings from each observer, and each image occupies a specific coordinate. (In this initial multi-dimensional space, for example, suppose that Stimulus #1 received a rating of 3 from Observer #1, but a rating of 5 from Observer #2. That would then fall on the coordinates of (3, 5, ...).) The PCA linearly combined the individual observer axes in order to find a set of transformed axes that capture the most variance. So if each observer's taste is fully independent of the others, then the stimuli would simply be spread out roughly to the same extent along any direction—and as a result, PCA will extract many different axes, with each one explaining only a relatively small proportion of the overall variance. However, if observers' tastes form some general patterns, then PCA will effectively be able to collapse those observers' dimensions, such that a smaller number of components can then explain more overall variance.

Supplemental Information

"Taste typicality" is a foundational

and multi-modal dimension

of ordinary aesthetic experience

Yi-Chia Chen, Andrew Chang, Monica D. Rosenberg, Derek Feng, Brian J. Scholl, and Laurel J. Trainor

Figure S1. Visual Stimuli, Related to STAR Methods. All 88 images used in the experiments are shown above, ranked by average ratings from the replication experiment. The best image is in the top left corner, and the images are ranked in descending order from left to right then top to bottom, with the worst image in the bottom right corner.