The Aesthetics of Mathematical Explanations

Samuel G. B. Johnson[1*] & Stefan Steinerberger[2*]
(sgbjohnson@gmail.com, stefan.steinerberger@yale.edu)
(*equal contributions)
[1]School of Management, University of Bath, Bath, BA2 7AY UK
[2]Department of Mathematics, Yale University, New Haven, CT 06520 USA

Abstract

Mathematicians often describe arguments as "beautiful" or "dull," and famous scientists have claimed that mathematical beauty is a guide toward the truth. Do laypeople, like mathematicians and scientists, perceive mathematics through an aesthetic lens? We show here that they do. Two studies asked people to rate the similarity of simple mathematical arguments to pieces of classical piano music (Study 1) or to landscape paintings (Study 2). In both cases, there was internal consensus about the pairings of arguments and artworks at greater than chance levels, particularly for visual art. There was also some evidence for correspondence to the aesthetic ratings of undergraduate mathematics students (Study 1) and of professional mathematicians (Studies 1 and 2).

Keywords: Psychology of mathematics; Explanation; Aesthetics; Reasoning; STEM education

Introduction

Explanations in science and mathematics are often imbued with aesthetic qualities. They can be "elegant" or "beautiful"; they can be "dull" or "trivial." Moreover, many scientists appear to adopt John Keats' maxim that "beauty is truth, truth beauty." Albert Einstein "was quite convinced that beauty was a guiding principle in the search for important results in theoretical physics" (Zee, 1999), while the physicist Paul Dirac (1963) even claimed that "it is more important to have beauty in one's equations than to have them fit experiment."

Is there a deep psychological reality underlying the perception that certain mathematical arguments are beautiful? One affirmative piece of evidence comes from a neuroimaging study (Zeki, Romaya, Benincasa, & Atiyah, 2014), in which professional mathematicians' brains were scanned while contemplating either "beautiful" or "ugly" mathematical equations. Individual judgments of mathematical beauty were correlated with activity in the same region of the medial orbito-frontal cortex that is known to track aesthetic judgments in other domains such as visual art and music (Ishizu & Zeki, 2011). This suggests that the underlying experience of mathematical beauty, at least for mathematicians, has a kinship with other forms of aesthetic experience.

However, it is plausible that these aesthetic experiences are unique to professional mathematicians. Indeed, in Zeki et al.'s (2014) study, the effort to scan layperson participants' brains was abandoned after an initial behavioral study failed to find *any* equations that participants found beautiful. Perhaps, then, the aesthetic experiences in mathematics are merely a by-product of the social practices of scientists and mathematicians.

Might it nonetheless be possible that novices share the aesthetic perception of career mathematicians? Previous research suggests that people do have intuitions about the acceptability of simple mathematical explanations (Johnson, Johnston, Koven, & Keil, 2017) and that such intuitions broadly track principles derived from philosophy of mathematics (e.g., Bolzano, 1817; Kitcher, 1975). But little is known about the intuitive *aesthetics* of such explanations, despite longstanding philosophical interest in the aesthetics (Lipton, 2004) and phenomenology (Gopnik, 1998) of explanation.

We report two behavioral studies of laypersons' aesthetic perception of mathematical arguments. Specifically, participants examined simple mathematical arguments and rated their similarity to various pieces of art—either music (in Study 1) or paintings (Study 2). This approach assumes that people can track broad aesthetic similarities across domains. Indeed, there is evidence for this broad capacity. Laypeople can detect which piece of music inspired which artwork (Ranjan, Gabora, & O'Connor, 2013), and students of painting can detect which of their peers created particular non-painting art (e.g., poetry; Gabora, O'Connor, & Ranjan, 2012). Our question is whether such abstract cross-domain aesthetic correspondences could also be detected for mathematics.

We seek answers to three broad questions: Are participants' responses internally consistent? Does education in higher mathematics alter this pattern? And to what extent are these layperson judgments consistent with those of professional mathematicians?

Study 1

Commentators have long noted an affinity between mathematics and music (Fauvel, Flood, & Wilson, 2006). Music has a mathematical structure, and the Greek mathematician Pythagoras worked out aspects of Western theory that persist to this day. Study 1 tests whether, in keeping with this affinity, the aesthetics of specific mathematical arguments intuitively correspond to different pieces of music.

Method

Participants were recruited from the online crowdsourcing platform Amazon Mechanical Turk (N = 299) and were from the United States. A subset of these participants (N = 90) had taken a university-level math course above the level of calculus, while a larger subset reported not having taken such coursework (N = 207). Given the lack of previous work on this topic, the sample size was chosen arbitrarily such that it was reasonably large (target N = 300); the same sample size was used in Study 2 to prevent experimenter degrees of freedom

Participants each read four mathematical arguments (see Table 1). For each argument, they were asked to first read and reflect on the argument. Then, on subsequent pages, participants rated the similarity of the argument to four different 20-second clips of classical music (see Table 2) on a scale from 0 ("not at all similar") to 10 ("very similar"). The arguments were presented in a random order, as were the musical clips, with each clip on a separate page.

Table 1: Arguments Used in Studies 1 and 2

Geometric	Sum of an infinite geometric series
Gauss	Gauss's summation trick for positive integers
Pigeonhole	Pigeonhole principle
Faulhaber	Geometric proof of a Faulhaber formula

Note. The exact text of these arguments can be found in Johnson and Steinerberger (2018).

Table 2: Musical Pieces Used in Study 1

Schubert	Moment Musical No. 4 (D 780), played by David Fray
Bach	Fugue from Toccata in E Minor (BWV 914), played by Glenn Gould
Beethoven	Diabelli Variations (Op. 120), played by Grigory Sokolov
Shostakovic h	Prelude in D-flat Major (Op. 87 No. 15), played by Adrian Brendle

Note. Participants listened to the first 20 seconds of each piece.

After the main task, a series of memory check questions was included to monitor whether participants had been attending to the materials. Participants were excluded from analysis if they incorrectly answered one-fourth or more of these questions (N = 73).

In addition to the Mechanical Turk sample, two additional sets of individuals were asked to complete the study for comparison. First, a sample of Yale undergraduates (N = 28) was recruited from Mathematics and Applied Mathematics courses, including Calculus, Linear Algebra, Abstract Algebra, and Probability Theory. Second, a sample of professional mathematicians (N = 4) was recruited from the first author's professional network. For these samples, participants were contacted by email and we included all participants who completed the principal measures.

Results and Discussion

The rankings of the pieces of music were not random: They reflected a degree of consensus which was shared to some extent across samples. This suggests that people have stable intuitions about the aesthetics of mathematics.

The mean ratings and consensus ranks assigned to each piece of music, separated by argument, are given in Table A1 (left side). These are computed for the Mechanical Turk sample as a whole, separately for the portions of the sample with and without higher mathematics coursework, for the sample of Yale undergraduates, and for the professional mathematicians. For statistical tests, we focus on the larger Mechanical Turk sample.

Did participants give similar ratings to one another? We tested this question in two ways. First, we calculated for each participant the correlation between that participant's ratings of the 16 items (i.e., 4 arguments x 4 pieces of music) and the mean ratings across the sample (leaving out that participant from the mean). These correlations were positive significantly more often than they were negative [145 out of 219 were positive; p < .001, sign test; 95% CI[.59, .72] on the proportion of positive correlations]. This was also true if we repeat the analysis separately on the participants who had taken higher mathematics [45 out of 68; p = .010; 95% CI[.54, .77]] and those who had not [93 out of 149; p < .001; 95% CI[.54, .70]] or if Spearman correlations are used instead.

A second way of testing this question is to compute Cronbach's alpha, treating each of the 16 ratings as an observation and each participant as a scale component. (This is similar to calculating the average correlation between each participant's scores and every other participant's scores.) These scores were fairly high (α = .72), indicating consistency across participants.

Did participants come to the same consensus as the experts? We tested this question too in two ways. First, similar to our above analysis, we computed for each participant the correlation between that participant's ratings of the 16 items and the mean rating for the other groups (expert mathematicians and students). The Mechanical Turk sample as a whole did not significantly agree with the mathematicians at greater than chance levels [120 out of 219 correlations were positive; p = .18; 95% CI[.48, .62]]. However, there was a marginally significant tendency for agreement among the higher-math sample [42 out of 68; p = .068; 95% CI[.49,.73]] but not the non-higher-math sample [76 out of 149, p = .87, 95% CI[.43,.59]]. The results were similar, albeit more statistically robust, for agreement with the student consensus. These correlations were positive significantly more often than chance for the sample as a whole [126 out of 219; p = .030, 95% CI[.51,.64], but this was driven principally by the higher-math subsample [42 out of 68; p = .068, 95% CI[.49,.73]] rather than the non-higher-math subsample [83 out of 149, p = .19, 95% CI[.47,.64]].

Second, we computed the correlations between the mean (consensus) ratings for the 16 items across different groups. First, there was a moderate but non-significant correlation between the higher-math and non-higher-math consensuses [$r(14)$ = .40, p = .12]. The Mechanical Turk sample as a whole was not significantly correlated with either the mathematicians [$r(14)$ = .06, p = .81] or the students [$r(14)$ = .40, p = .12]. Consistent with the above analysis, however, the correlations were numerically somewhat stronger for the higher-math subsample [$r(14)$ = .27, p = .31 with mathematicians and $r(14)$ = .58, p = .012 with students], with the student correlation reaching statistical significance. These correlations were all nonsignificant for the non-higher-math subsample [$r(14)$ = –.09, p = .75 and $r(14)$ = .21, p = .43].

Discussion. Overall, these results suggest that laypeople have intuitions about the aesthetics of mathematical arguments, which were internally consistent within the sample of laypeople. For those who had taken higher mathematics (but not those who had not), there was some consensus with undergraduate math students and possibly with professional mathematicians. This overall pattern suggests that although even laypeople have intuitions about the aesthetics of mathematical arguments, these intuitions may sharpen with math instruction.

Study 2

Study 2 tested whether the aesthetic structure of mathematical arguments is limited to music—which has long invited mathematical comparisons—or whether people can also perceive similarities in the aesthetics of mathematics with other artistic mediums. Here, we used landscape paintings as an aesthetic medium that is not frequently imbued with mathematical character.

Method

Participants were recruited from Mechanical Turk (N = 300). Similar to Study 1, a subset of these participants (N = 99) had taken a higher mathematics course, while most participants had not (N = 201). Participants were excluded if they failed the same check questions used in Study 1 (N = 67) or did not produce a complete set of ratings (N = 1).

The procedure was identical to Study 1, except participants rated the similarity of each argument to four different landscape paintings (see Table 3).

In addition to the Mechanical Turk sample, a sample of professional mathematicians (N = 8) was recruited from the first author's professional network.

Table 3: Paintings Used in Study 2

Yosemite	Looking Down Yosemite Valley, California (Albert Bierstadt)
Rockies	A Storm in the Rocky Mountains, Mt. Rosalie (Albert Bierstadt)
Suffolk	The Hay Wain (John Constable)
Andes	The Heart of the Andes (Frederic Edwin Church)

Results and Discussion

The similarity ratings were lower overall in Study 2 than in Study 1, consistent with the idea that music is imbued with a more mathematical character. However, the associations between different artworks and different arguments followed a consistent pattern, indeed to a greater degree than in Study 1. The mean ratings and consensus ranks are given in Table A1 (right side).

Using the same method as Study 1, we calculated the correlations between each participant's ratings and the sample consensus (leaving that participant out). As in Study 1, these correlations were more often positive than negative [156 out of 211; p < .001, sign test; 95% CI[.67, .80]]. Alpha was even higher than in Study 1 [α = .93], indicating strong consistency across participants.

This consensus was also consistent across groups of participants. Using the same method as Study 1, participants' individual ratings in the Mechanical Turk sample were correlated with above-chance frequency with the mean ratings of the professional mathematicians [135 out of 211 correlations were positive; p < .001; 95% CI[.57, .70]]. Further, the mean ratings across groups were very similar. The subsamples with and without higher mathematics training were correlated at a remarkable $r(14)$ = .94 [p < .001], while the Mechanical Turk group as a whole correlated significantly with the mathematicians [$r(14)$ = .51, p = .046].

Discussion. Mathematical arguments are more often associated with musical rather than visual aesthetics, yet participants were even more able to associate arguments with landscape paintings in Study 2 than to associate them with classical music in Study 1. This suggests that there is a deep aesthetic structure to mathematics that even laypeople can perceive, and which is not limited to comparisons with a single aesthetic medium.

General Discussion

Mathematicians have strong intuitions about the beauty of mathematics. Our studies suggest that laypeople too have intuitions about the aesthetics of arguments. Participants' ratings of how mathematical arguments corresponded to different pieces of classical music internally consistent and correlated somewhat with expert judgments (Study

1), while such ratings were even more consistent for visual art (Study 2). These internally reliable judgments complement recent findings in other aesthetic domains. For instance, people share aesthetic knowledge within a culture for geometric figures (Westphal-Fitch & Fitch, 2017) and people can form stable aesthetic judgments about paintings on time-scales as short as 50ms (Verhavert, Wagemans, & Augustin, 2018).

One general concern about these results is that we are conflating reliability with validity. That is, we have shown that participants' judgments are internally consistent, but have not shown that these are driven by the "objective" aesthetics of the arguments. While we recognize this concern, several observations ameliorate this problem. First, we can treat the ratings of expert mathematicians as most closely approaching the aesthetic ground truth. When we do so, we find either that both participant groups track these expert judgments (in Study 2) or that only the more experienced participants track the expert consensus (in Study 1). These are the sort of results one would expect if lay participants are tapping into the same underlying truth as the experts. Second, we can exploit differences in the consistency of aesthetic intuitions themselves, which are known to be stronger for less abstract stimuli (Vessel & Rubin, 2010). Correspondences were indeed stronger between mathematical arguments and (less abstract) paintings, compared to (more abstract) music—even though the similarity ratings themselves were higher for music.

A related concern is that participants' similarity judgments may have been based on superficial rather than deep (e.g., aesthetic) similarities. For example, two proofs involved pictures of squares (Geometric and Faulhaber), so perhaps participants could have matched both proofs to paintings including more right angles. The data do not bear out this specific example (lay participants favored different paintings for these two proofs), but we could not possibly enumerate, much less eliminate, all such possibilities. That said, several findings suggest this is unlikely to be the main driver. First, very few of the participants' debriefing comments could be interpreted as supporting such superficial strategies (see below for more discussion about participants' reported strategies). Second, one would expect mathematicians' judgments to be relatively less contaminated by such superficial similarities, yet laypersons' judgments tended to be similar to the professionals'. Finally, although such superficial similarities are plausible for visual art, it is harder to generate such accounts for pieces of classical music, which is itself highly abstract.

These results inform longstanding issues in the philosophy of explanation and mathematics. Some have argued that people infer that an explanation is true when it strikes them as elegant, beautiful, or satisfying—that is, aesthetically pleasing (Johnson, 2017; Lipton, 2004). Previous studies do suggest that people find explanations likely to the extent that they are satisfying (Khemlani, Sussman, & Oppenheimer, 2011; Lombrozo, 2007) and that satisfying explanations are often "truth-tracking" in the sense that they conform to the laws of probability (Johnson, Rajeev-Kumar, & Keil, 2016; Johnson, Valenti, & Keil, 2017; Johnston et al., 2016). However, the current studies are the first, to our knowledge, to directly demonstrate an aesthetic component to lay explanation. Future research can build on these results in several ways.

First, a key question is whether these results are cross-culturally universal. In some cases, aesthetic preferences do appear to be similar across cultures (e.g., Palmer et al., 2013). However, given evidence that individuals from Western and East Asian cultures differ in their approach to logical reasoning (e.g., Peng & Nisbett, 1999), might people from different cultures have different aesthetic sensibilities about mathematical arguments?

Second, what is the relationship between aesthetic preferences and the ability of children to evaluate mathematical arguments? Given that other forms of explanatory reasoning seem to arise by age 4 (Bonawitz & Lombrozo, 2012; Johnston et al., 2017), might aesthetic preferences underlie the development of these broader capacities? If they are early-emerging, can aesthetic preferences be harnessed to facilitate math education?

Third, what cues are people relying on to form these aesthetic judgments? For example, beauty is strongly associated with pleasure (Brielmann & Pelli, 2017), and aesthetic experiences are often mediated by emotion more generally (Palmer et al., 2013). In other domains, such as photography, semantic content is known to be an important driver of aesthetic preferences (Vessel & Rubin, 2010). Eliciting more multi-dimensional ratings from participants (e.g., Blijlevens et al., 2017) would help to uncover the basis for the correspondences.

One intriguing possibility is that participants were relying on "collative properties" (e.g., Cupchik & Berlyne, 1979; Marin & Leder, 2013). These are broad aesthetic properties about the organization of a stimulus—such as order, complexity, and ambiguity—which generate affective reactions and generalize across different modalities, such as music and visual art (and perhaps mathematical arguments). Many participants cited such dimensions when asked to describe their strategies, particularly invoking complexity and clarity as useful dimensions that drove cross-domains judgments.

Finally, given the parallels between laypeople and mathematicians, would laypeople also rely on the same brain regions for appreciating mathematical arguments as for appreciating other forms of beauty, as mathematicians do (Zeki et al., 2014)? Might there even be *specific* neural correspondences between particular artworks and arguments, as we found here for similarity ratings?

While mathematics is important to society because it is useful, some have argued that it is the beauty of mathematics that justifies its importance as a field of study (Hardy, 1940). Indeed, these results suggest, the beauty of mathematics may be deeply human.

Acknowledgements

We thank Woo-kyoung Ahn for discussion, and Asher Auel, Miki Havlickova, John Hall, Brett Smith, and Sarah Vigliotta for assistance with data collection. Funding was provided by a Dissertation Research Award to the second author from the American Psychological Association.

References

Blijlevens, J., Thurgood, C., Hekkert, P., Chen, L.L., Leder, H., & Whitfield, T.W.A. (2017). The Aesthetic Pleasure in Design Scale: The development of a scale to measure aesthetic pleasure for designed artifacts. *Psychology of Aesthetics, Creativity, and the Arts, 11*, 86–98.

Bolzano, B. (1817). *Purely analytic proof of the theorem that between any two values which give results of opposite sign there lies at last one real root of the equation*. (Trans. S. Russ.)

Bonawitz, E.B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental Psychology*, *48*, 1156–1164.

Brielmann, A.A., & Pelli, D.G. (2017). Beauty requires thought. *Current Biology*, *27*, 1506–1513.

Cupchik, G.C., & Berlyne, D.E. (1979). The perception of collative properties in visual stimuli. *Scandinavian Journal of Psychology*, *20*, 93–104.

Dirac, P.A.M. (1963). The evolution of the physicist's picture of nature. *Scientific American*, *208*, 45–53.

Fauvel, J., Flood, R., & Wilson, R. (2006). *Music and mathematics: From Pythagoras to fractals*. New York, NY: Oxford University Press.

Gabora, L., O'Connor, B. P., & Ranjan, A. (2012). The recognizability of individual creative styles within and across domains. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 351-360.

Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines*, *8*, 101–118.

Hardy, G.H. (2004). *A mathematician's apology.* Cambridge, UK: Cambridge University Press. (Original work published 1940).

Ishizu, T., & Zeki, S. (2011). Toward a brain-based theory of beauty. *PLosS ONE*, *6*, e21852.

Johnson, S.G.B. (2017). *Cognition as sense-making* (Unpublished doctoral dissertation). Yale University, New Haven, CT.

Johnson, S.G.B., Johnston, A.M., Koven, M.L., & Keil, F.C. (2017). Principles used to evaluate mathematical explanations. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 612–617). Austin, TX: Cognitive Science Society.

Johnson, S.G.B., Rajeev-Kumar, G., & Keil, F.C. (2016). Sense-making under ignorance. *Cognitive Psychology*, *89*, 39–70.

Johnson, S.G.B., & Steinerberger, S. (2018). The aesthetic psychology of mathematics. Available at Social Science Research Network (SSRN): https://papers.ssrn.com/abstract=3062208.

Johnson, S.G.B., Valenti, J.J., & Keil, F.C. (2017). Simplicity and complexity preferences in explanation: An opponent heuristic account. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 606–611). Austin, TX: Cognitive Science Society.

Johnston, A.M., Johnson, S.G.B., Koven, M.L., & Keil, F.C. (2017). Little Bayesians or little Einsteins? Probability and explanatory virtue in children's inferences. *Developmental Science, 20, e12483.*

Khemlani, S.S., Sussman, A.B., & Oppenheimer, D.M. (2011). *Harry Potter* and the sorcerer's scope: Latent scope biases in explanatory reasoning. *Memory & Cognition*, *39*, 527–535.

Kitcher, P. (1975). Bolzano's ideal of algebraic analysis. *Studies in the History and Philosophy of Science*, *6*, 229–269.

Lipton, P. (2004). *Inference to the best explanation* (2nd Edition.). London, UK: Routledge.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*, 232 –257.

Marin, M.M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings, and music. *PLoS One*, *8*, e72412.

Palmer, S.E., Schloss, K.B., Xu, Z., & Prado-Léon, L.R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, *110*, 8836–8841.

Peng, K., & Nisbett, R.E. Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*, 741–754.

Ranjan, A., Gabora, L., & O'Connor, B. (2013). The cross-domain re-interpretation of artistic ideas. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 3251–3256). Austin, TX: Cognitive Science Society.

Verhavert S., Wagemans J., Augustin M. (2018). Beauty in the blink of an eye: The time course of aesthetic experiences. *British Journal of Psychology*, *109*, 63–84.

Vessel, E.A., & Rubin, N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. *Journal of Vision*, *10*.

Westphal-Fitch, G., & Fitch, W.T. (2010). Beauty for the eye of the beholder: Plane pattern perception and production. *Psychology of Aesthetics, Creativity, and the Arts*, *11*, 451–456.

Zee, A. (1999). *Fearful symmetry: The search for beauty in modern physics.* Princeton, NJ: Princeton University Press.

Zeki, S., Romaya, J.P., Benincasa, D.M.T., & Atiyah, M.F. (2014). The experience of mathematical beauty and its neural correlates. *Frontiers in Human Neuroscience*, *8*, 68.

Appendix

Table A1: Results of Studies 1 and 2

	Schubert	Bach	Beethoven	Shostakovich	Yosemite	Rockies	Suffolk	Andes
	Mechanical Turk Sample (All)							
Geometric	**4.76 (1)**	4.39 (3)	4.36 (4)	4.62 (2)	**3.51 (1)**	2.99 (4)	3.30 (2)	3.05 (3)
Gauss	4.61 (3)	**5.11 (1)**	4.67 (2)	4.31 (4)	2.38 (2)	2.23 (3)	**2.43 (1)**	1.96 (4)
Pigeonhole	4.52 (3)	4.42 (4)	**4.89 (1)**	**4.83 (2)**	**2.42 (2)**	2.21 (4)	2.25 (3)	**2.49 (1)**
Faulhaber	4.32 (4)	4.59 (3)	**5.04 (2)**	**5.06 (1)**	2.97 (2)	2.75 (3)	**3.21 (1)**	2.44 (4)
	Mechanical Turk Sample (No Higher Math Background)							
Geometric	4.66 (2)	4.38 (4)	4.49 (3)	**4.79 (1)**	**3.41 (1)**	2.88 (4)	3.08 (2)	2.88 (3)
Gauss	4.47 (3)	**5.01 (1)**	4.63 (2)	4.29 (4)	**2.31 (1)**	2.15 (3)	**2.28 (2)**	1.88 (4)
Pigeonhole	4.33 (4)	4.35 (3)	**4.98 (1)**	**4.96 (2)**	2.23 (2)	1.99 (4)	2.09 (3)	**2.37 (1)**
Faulhaber	4.53 (4)	4.58 (3)	5.07 (2)	**5.19 (1)**	2.77 (2)	2.64 (3)	**3.06 (1)**	2.35 (4)
	Mechanical Turk Sample (Higher Math Background)							
Geometric	**5.04 (1)**	4.43 (2)	4.11 (4)	4.30 (3)	**3.76 (2)**	3.27 (4)	**3.79 (1)**	3.44 (3)
Gauss	4.91 (2)	**5.32 (1)**	4.71 (3)	4.32 (4)	2.55 (2)	2.40 (3)	**2.77 (1)**	2.13 (4)
Pigeonhole	**4.97 (1)**	4.60 (3)	4.73 (2)	4.58 (4)	**2.85 (1)**	2.71 (3)	2.62 (4)	**2.75 (2)**
Faulhaber	3.91 (4)	4.61 (3)	**5.06 (1)**	4.78 (2)	3.45 (2)	3.01 (3)	**3.56 (1)**	2.66 (4)
	Students							
Geometric	5.17 (2)	**5.86 (1)**	4.72 (4)	4.78 (3)				
Gauss	5.21 (2)	**5.51 (1)**	4.76 (4)	5.20 (3)				
Pigeonhole	**4.99 (2)**	4.96 (3)	**5.09 (1)**	4.76 (4)				
Faulhaber	4.04 (4)	4.34 (3)	**6.01 (1)**	5.11 (2)				
	Professional Mathematicians							
Geometric	5.30 (2)	**5.80 (1)**	3.50 (3)	2.62 (4)	**4.19 (1)**	3.30 (2)	3.20 (3)	2.17 (4)
Gauss	3.38 (4)	**7.75 (1)**	3.95 (3)	5.72 (2)	**3.40 (2)**	**3.46 (1)**	2.84 (3)	2.34 (4)
Pigeonhole	**6.80 (1)**	5.20 (2)	4.88 (3)	4.47 (4)	2.27 (2)	**2.59 (1)**	2.31 (3)	2.08 (4)
Faulhaber	5.95 (2)	**6.28 (1)**	3.75 (4)	5.62 (3)	**4.19 (1)**	3.61 (2)	2.95 (3)	2.81 (4)

Note. Entries are the mean ratings assigned to each artwork, with consensus ranks for each argument given in parentheses. Entries are bolded if they are the highest ranked piece or 0.1 or fewer scale points away from the highest rank.