Soklynin Nou

11/17/22

A Reevaluation of Birkhoff's Formula on Polygons

Abstract

George David Birkhoff attempted to analyze the elements that contribute to an object's beauty, *Order* and *Complexity*, in his book *Aesthetic Measure*. He then proposed a formula to objectively evaluate a polygon's objective beauty. This paper proposes a modified formula to match the predicted values of Birkhoff's formula and the survey results. The survey sampled 91 students from four different math-related classes at the College of San Mateo. The two data sets of Birkhoff's formula and the survey results are then compared and contrasted to find coefficients for each variable that make Birkhoff's formula align more with the survey results. Calculating the coefficient of determination for each of these sets based purely on the results of Birkhoff's formula is 0.1441 for the first set and 0.3803 for the second set, which means that the independent variables of Birkhoff's formula does very little to predict the outcome of the survey conducted. With the newly added variables, the formula now produces results that more closely align with the survey conducted.

<u>Introduction</u>

George Birkhoff has attempted to bridge the gap between Mathematics and art by looking at the aesthetic of an object from a mathematical standpoint. George David Birkhoff was an American mathematician born in 1884 and was one of the most important leaders in American mathematics of his generation. Even though he is a mathematician, he has always had an interest in the aesthetics of different forms of art; he is interested in what makes arts such as music, poems, sculpture, vases, and painting appealing to humans. Birkhoff had written many books such as *Relativity and Modern Physics* and *Dynamical Systems*, which are mostly about mathematical theories. But, In 1933, Birkhoff published a book on the matter of evaluating aesthetics called *Aesthetics Measure*, which was published in the *Harvard University Press*. In the book, he analyzes a lot of the different forms of art using a formula he created. . In the case of polygons, Birkhoff defines *O* (order) as the sum of *V* (Verticle symmetry), *E* (Equilibrium), *R* (Rotational Symmetry), and *HV* (the relation of the polygon to a horizontal-verticle grid) subtracted by *F*, which is a list of factors defined by Birkhoff that makes the polygon unsatisfactory. *C* (complexity) represents the number of lines all edges of the polygon lie on. This is not to be confused with the number of sides the polygon has as two edges could lie on one line. All these variables can only be discrete, whole numbers. In the end, Birkhoff's formula for polygons is $A(S) \ = \frac{V+E+R+HV-F}{c}$. This is the particular formula the paper will be using.

This formula was studied in many other papers such as Veronika Douchová's review and critical evaluation of George D. Birkhoff's work as well as her discussions on the influence it has on further research in the field. The paper also mostly focuses on polygons. But, much of it is just applying the formula verbatim. Others like Hans Jürgen Eysenck, a German-British Psychologist, partially inspired this paper. He took a larger-scale survey and discussed the correlation between the results and the output of Birkhoff's formula. He displayed his results with other similar studies, which showed that there are a lot of variations in the correlation with Birkhoff's formula. The correlation coefficient ranges from 0.7 to 0.05.

This paper will expand upon the formula by making it better reflect the survey results that we conducted. Birkhoff's formula assumes that people value each of the different factors of a shape equally. This paper as well as other papers such as Hans Jürgen Eysenck's paper shows the results of Birkhoff's formula and the surveys have very little correlation with each other. To achieve a more accurate result, we give the independent variable of Birkhoff's formula varying importance by attracting a coefficient of to each. These coefficients could be considered as how much people like that specific attribute of a polygon. There are assumptions made in the paper and justifications for those assumptions. This will be further discussed in later sections of the paper.

The use of math in analyzing art may seem at first glance to be a reductive approach, stripping away the emotional and subjective elements that make art such a powerful and enduring human activity. However, math can actually provide a way of understanding and appreciating the underlying structure and patterns in art that may not be immediately apparent to the naked eye. By using mathematical principles such as symmetry, proportion, and geometry, we can gain a deeper understanding of how an artist has created their work and the effects they are trying to achieve. In this way, math can become a tool for enhancing our appreciation and understanding of art, rather than diminishing it.

New Formula

As previously mentioned, the formula this paper will analyze is Birkhoff's formula on the polygon. Birkhoff identifies five components that contribute to the order of an object. This paper will assume that these are the only factors that contribute to the order of a polygon. Components such as the symbolism the polygon represents are not considered just like Birkhoff. As James R. Newman noted in a book titled *In The World of Mathematics*, Birkhoff doesn't factor in associations to the polygon, which is the reason why the Greek cross has a relatively high value of 0.75 while the Roman cross only has a value of 0.25. This means that when finding coefficients for each variable of the formula, this paper will not consider any other factors that might contribute to the aesthetic measure of an object. Not considering any other associations is also relevant when a survey is conducted for this paper. The coefficients represent how much each component of the function is worth relative to the other components. This means that the values of each coefficient are arbitrary if isolated on their own. For example, if the coefficient for vertical symmetry is $x_1 = 0.5$ and the coefficient for rotational symmetry is $x_2 = 0.38$, we could then conclude that vertical symmetry is considered 31.57% more valuable than rotational symmetry. Therefore, the numerator of the new modified function becomes, $O = x_1V + x_2E + x_3R + x_4HV - F$, where x_i is the coefficient of each variable. There is also no denominator. This will be further explained in the “Multiple Linear Regression” section. With these new elements, the modified function now comes to:

$$A(S) = x_1V + x_2E + x_3R + x_4HV - F$$

Where *A* is the aesthetic measure function and *S* is the specific shape being analyzed. The last variable also doesn’t have a coefficient. This will be explained in the “Calculations” section.

For the independent variables given, if vertical symmetry exists, **V** = 1. If not, then **V** = 0. Birkhoff defines equilibrium (**E**) as the relative position of the center of mass of the object. This can be thought of as whether or not the object is balanced, assuming the object’s density is constant throughout. **E** could take the value of 1, 0, or -1. **E** = 1 when the center of mass of the object lies between the vertical lines crossing the extremal points of the shape, and the distance to these lines is at least 1/6 of the width of that object. **E** = 0 if the center of mass lies between the vertical lines but at a smaller distance. Otherwise, **E** = -1. The rotational symmetry (**R**) of an object is the number of rotations possible before the shape returns to its original position. The number of these rotations is denoted by q. **R** is then $\frac{q}{2}$ if $q \leq 6$, 3 if $q > 6$, 1 for other cases if rotational symmetry exists and q is even, and 0 otherwise. **HV** is the shape's relation to a horizontal and vertical grid. The value could be 2, 1, or 0. If all the edges lie in parallel with a grid, then the value is 2. If the edges of the shape form lines with the same angle and all edges of the polygon lie on that grid, then the value is 1. Otherwise, we set **HV** = 1. **F** is what Birkhoff considers to be detrimental towards the aesthetics of the object. Veronika Douchová has summarized the parameters Birkhoff outlined in his book. She wrote that **F** = 0 if all these conditions are met:

1. there are no distances between vertices, edges and between a vertex and an edge which are too small. A distance is too small if it is less than 1/10 of the maximal distance between the vertices of the polygon.
2. The angles between the edges are not too small (set as less than 20°).
3. There are no irregularities. This is defined that no movement of a vertex by a small amount (less than 1/10 of the distance to the nearest vertex) may impact the values of **V**, **R** or **HV**.
4. There are no projecting edges.
5. There is at most one type of concave angle.
6. If we view the vertical and horizontal directions as one direction, then the polygon cannot have more than two types of direction.
7. There is a symmetry which prevents both **V** and **R** from taking the value 0.

If one of these conditions is not met, then **F** = 1. If two or more are not met, then **F** = 2 (Douchová). Some of the parameters Birkhoff outlined could be considered arbitrary, but the paper assumes that these parameters are correct.

Correlations with Survey

A survey has been conducted for the paper to find the correlation between the results and the values achieved from Birkhoff's measure. The results of the survey are ranked data whereas Birkhoff calculations produced “objective” value.

But, Birkhoff’s aesthetic value is actually an arbitrary value that is only useful when compared to another shape’s value. Because of this, we can create new values that could be considered a bit arbitrary in isolation. The values are created by taking the mean ranking of an individual shape and subtracting it from 10. So, a mean ranking of 2.3 will produce a value of 7.7. This approach is sensible because there is no aesthetic measure threshold to make it have other properties. This is in contrast to measurements of newtons, where a certain value can be the difference between a stone breaking or not. But, these values are not completely arbitrary as the closer the value is to nine, the better it is. These new values are then put into a program that computes the best-fitting coefficient for each of the components based on the output received from each input of the function.

A way to test the correlation between the two results is by calculating the r-value. The r value represents the correlation coefficient, which is a statistical measurement of the strength of a linear relationship between two variables. The correlation coefficient can be calculated using the equation shown below:

$$r = \frac{\sum_{i=1}^{n}[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Where r is the Pearson Correlation Coefficient

x_i is the x variable sample

y_i is the y variable sample

$\bar{x}$ is mean value of x

$\bar{y}$ is mean value of y

n is the number of shapes

The output of the formula will be in the interval $-1 \leq r \leq 1$ and the closer r is to ± 1, the more the outputs of the x and y values are correlated. This implies that a coefficient of determination, or the square of the correlation coefficient (r^2), of 1 means there is a perfect correlation between the two variables. If $r = \pm 0.7$, then $r^2 = 0.49$, which is about halfway between 1. This can be interpreted as 49% of the independent variable is responsible for the dependent variable. Using this test, we could see that comparing Birkhoff's measure to the survey results produces a coefficient of determination of around 0.1441 for the first set and 0.3803 for the second. These results show that there is very little correlation between Birkhoff's aesthetic measure and the survey conducted.

Multiple Linear Regression

Multiple linear regression will be used to find the coefficients for each element of the function as it can test the coefficients using all the shapes. This method finds a way to get the various independent variables and various dependent variables to be linearly related to each other. Excel was the instrument used to compute the coefficients of these variables. The formula used to calculate the coefficients using the formula below:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_n x_{in} + \epsilon$$

Where y_i is the dependent variable

x_{in} is the independent variable

β_0 is the y-intercept

β_n is the coefficient

ϵ is the residual

The modified formula is similar to the equation for multidimensional linear regression and so it is the main formula that needs to be computed. The output will not be the same as the observed data but it could be very close. This difference between the observed data and the predicted data is the residual.

We only modified *Order* coefficients ($O = x_1V + x_2E + x_3R + x_4HV + F$). The reason why we don't add a coefficient for complexity, therefore assuming that the weight of complexity is one, is because it makes the formula above non-linear. Non-linear models are currently beyond the author's field expertise.

Interpretation of Survey Results

The raw ranked data of the survey is then used to find the mean of each shape. This mean is treated as the aesthetic value of the object when subtracted from 10. The number the mean is subtracted from can be arbitrary because the only change would be in the y intercept. A greater value makes the y intercept higher and a lower one makes the y intercept lower. The y Intercept of the multiple linear regression is not relevant to the modified formula.

We can test to see if the mean of the survey data accurately represents what people picked. A mean of 3 could mean that half of the sample picks 1 and the other half pick 5 or everyone picks 3. Having a mean of 3 doesn't accurately represent a data set 1s and 5s as these numbers massively deviate from the mean. To test the accuracy of the mean, the formula for standard deviation is used:

$$SD = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}}$$

Where SD is the Standard deviation

x_i is each value of the sample

μ is the sample mean

N is the sample size

The standard deviation (SD) and mean for each shape is shown below:

	SD=1.49	Mean= 2.90		SD=1.12	Mean= 2.4
	SD=1.18	Mean= 2.77		SD=1.45	Mean= 3.87
	SD=1.21	Mean= 3.32		SD=1.28	Mean= 2.68
	SD=1.46	Mean= 2.38		SD=1.10	Mean= 2.73
	SD=1.40	Mean= 3.62		SD=1.48	Mean= 3.49

Considering that the sample is of discrete values in a small range, the standard deviation of many of these shapes is relatively small.

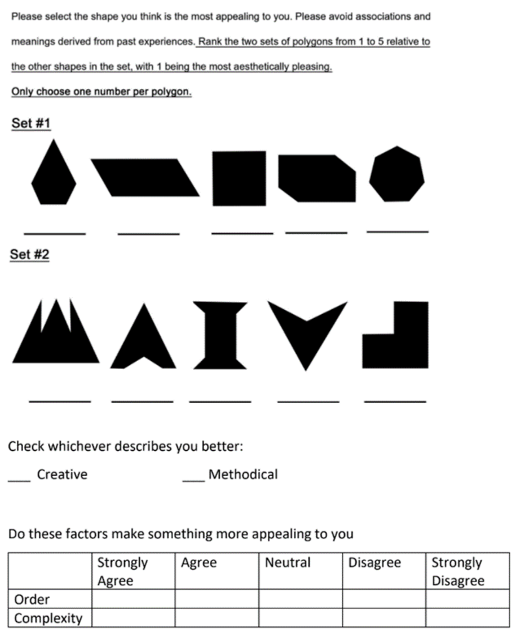
Methodology

This section of the paper highlights the methods used to properly evaluate the relevancy of Birkhoff's formula for polygons, both around the release of *Aesthetic Measures* and the modern day. The data used comes in the form of surveys conducted at both of the times mentioned above.

Hans Eysenck has compiled many investigations, some of which he has done himself, on the correlations of Birkhoff's formula on polygons with various opinions from people of varying knowledge on the subject. For Eysenck's own research, he had fourteen observers rank two sets of 32 polygons. The observers consist of seven men and seven women, some of whom are artists, students, professional men and women, teachers, stenographers, and psychologists. But, none of them was familiar with Birkhoff's theories to eliminate biases. He found that the correlation between his result of the survey, including an additional one he did with 12 participants, with Birkhoff's formula ranges from 0.48 to 0.62 using the correlation coefficient (Eysenck). This data can also be compared to a smaller survey that is being conducted specifically for this paper.

The survey being conducted consists of two sections. The first section asks the participants to rank two sets of five polygons relative to the other polygons in the set. The decision to break the polygons into two sets and have only a few in each set avoids the participants from overthinking their choices, which makes their decision-making worse when presented with too many options. The phenomenon is highlighted in a book called *The Paradox of Choice* by American psychologist Barry Schwartz. The order of each polygon's position is also randomized to remove its position from being a potential for bias. But, due to practical reasons, there are only three versions of the survey. This is because having it be truly unique requires 120 versions (5!) for each set, most of which only have one difference. The size of each polygon is also quite similar so that shape is the only variable. The second section of the survey asks two questions. The first asks the participants whether they consider themselves more creative or analytical. The second question asks them to rate the importance of order and complexity on a 5-point scale in making an object more appealing to them. The survey will be given to at least two classes of mathematics classes, one of which is a liberal art math class. A few more participants from an architecture class have also been given the survey to act as a potentially more creative group. The survey samples taken could potentially lead to a bias in the result because it only consists of college students in a single college. This means that the designs around the area could have an influence on the participants. Even though the survey asks the participants to avoid associations and meanings derived from past experiences, it is still possible that there are subconscious associations happening. The sample of college students taken also isn't that diverse. The vast majority of the participants will be from math classes. So, there might be little representation of the creative perspective. An unknown variable in the survey is that we don't have information about the students' majors; we only know one class that they take.

A sample of this survey can be seen below:



Please select the shape you think is the most appealing to you. Please avoid associations and meanings derived from past experiences. Rank the two sets of polygons from 1 to 5 relative to the other shapes in the set, with 1 being the most aesthetically pleasing.

Only choose one number per polygon.

Set #1

_____ _____ _____ _____ _____

Set #2

_____ _____ _____ _____ _____

Check whichever describes you better:

___ Creative ___ Methodical

Do these factors make something more appealing to you

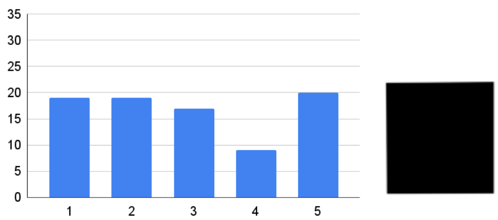
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Order					
Complexity					

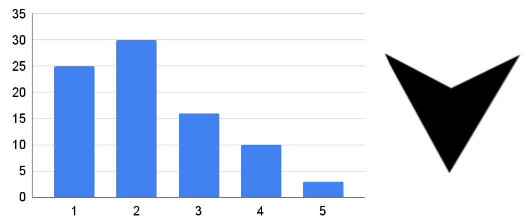
Two extra sections were also added as they might be useful but the data they produced was not used in the end. One of them consisted of a small personality test to see whether the participants thought of themselves as creative or methodical. Another is to see what the participants think makes an object more appealing to them. The raw data is in this google sheets: Survey Response (Paper)

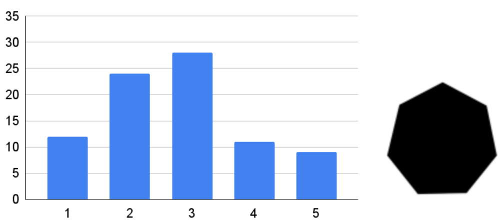
The survey by Professor Jacqueline Yahn was a trial run for the survey and was not considered.

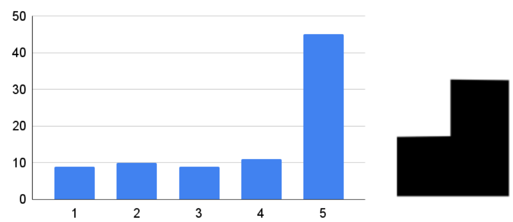
Survey Results

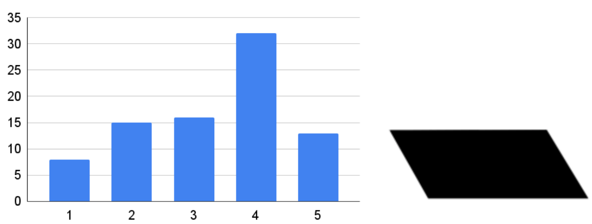
84 responses were counted and the results of the survey show that there is a lot of agreement between people ranking the shapes. The Y-axis represents the number of people that picked a ranking. All the bar graphs have a Y-axis from 0 to 35 except for the second shape of the second set. This shape's Y axis is extended because 45 people agreed on one ranking.

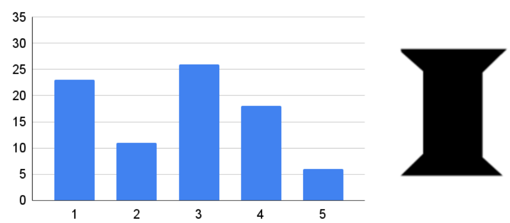


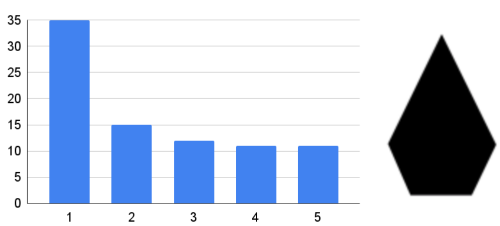


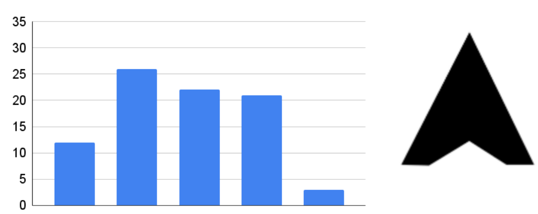


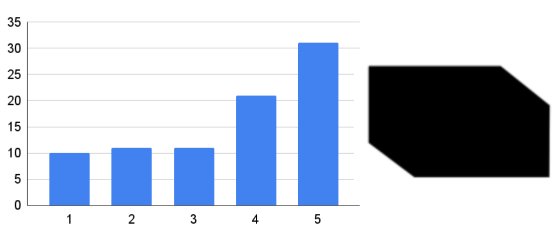


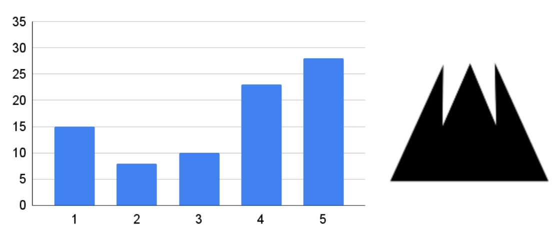












As we can see in the bar graph, most of the shapes show spikes in a specific ranking. This means that the data is very unlikely to be random. The fourth shape of the first set and the second shape of the second set is the best example of a clear preference. Although, there are still shapes where the data could be considered random. The first shape of the first set is one such example as the response seems to be almost equally distributed.

Calculations

The data analysis tool in Excel was used to compute the coefficients as well as many other values relevant to the study.

First Set (With Complexity)

Regression Statistics			*Coefficients*	Observation	Predicted Y	Residuals
Multiple R	0.857569099	Intercept	37.65	1	28.4	1.06581E-14
R Square	0.735424759	X Variable 1	0	2	50.61	7.10543E-15
Adjusted R Square	-1.058300964	X Variable 2	0.45	3	32.8	-7.28
Standard Error	10.29547473	X Variable 3	4.17	4	38.1	7.10543E-15
Observations	5	X Variable 4	-9.02	5	32.8	7.28

Second Set (With Complexity)

Regression Statistics			*Coefficients*	*Observation*	*Predicted Y*	*Residuals*
Multiple R	0.997790745	Intercept	50.335	1	31.04	0
R Square	0.99558637	X Variable 1	0	2	36.78	0
Adjusted R Square	-0.017654518	X Variable 2	-5.74	3	58.56	0
Standard Error	1.378858223	X Variable 3	13.965	4	44.595	-0.975
Observations	5	X Variable 4	-13.555	5	44.595	0.975

First Set (Without Complexity)

Regression Statistics			*Coefficients*	*Observation*	*Predicted Y*	*Residuals*
Multiple R	0.975614809	Intercept	6.79	1	7.1	0
R Square	0.951824255	X Variable 1	0	2	7.23	0
Adjusted R Square	-0.19270298	X Variable 2	0.83	3	6.53	-0.15
Standard Error	0.212132034	X Variable 3	-0.13	4	7.62	0
Observations	5	X Variable 4	-0.13	5	6.53	0.15

Second Set (Without Complexity)

Regression Statistics			*Coefficients*	*Observation*	*Predicted Y*	*Residuals*
Multiple R	0.913761518	Intercept	5.26	1	7.76	-8.88178E-16
R Square	0.834960112	X Variable 1	0	2	6.13	-8.88178E-16
Adjusted R Square	-0.660159554	X Variable 2	1.63	3	7.32	-8.88178E-16
Standard Error	0.537401154	X Variable 3	0.43	4	6.89	0.38
Observations	5	X Variable 4	0.87	5	6.89	-0.38

There are two types of calculations in the figure above: With and without complexity. The coefficient of one set is tested on the other sets for both types of results. An oversight the paper made is that multiple linear regression cannot calculate results if the number of independent variables in one function is equal to the number of observed outputs of those functions. This means that one variable of Birkhoff's formula has been left out. The variable chosen is *F*. This is because the values for *F* in the first set are all zeroes and the second set only has one zero value while the rest is -2. With this taken into consideration, the squared of the r-value is roughly 0.95 for the first set and 0.83 for the second set when complexity is not factored in. This means that the coefficients of both sets a lot more accurately describe the survey results than Birkhoff’s formula. The r-squared value is roughly 0.73 for the first set and 0.99 for the second set when complexity is factored in. This flip in the amount of correlation could possibly be because the second set is more complex. So, factoring in complex better reflects the output of the second set. When applying the coefficients of the first set to the second, we could see that the R-squared value for the second set becomes 0.4165, which is better than having no coefficients. It is also the best result since all the other combinations of coefficients produced r-squared values less than this.

<u>Conclusion</u>

With Birkhoff’s formula not reflecting the results of multiple surveys, coefficients have been added to the independent variables. These coefficients show what factors people prefer and which people think are less important. The results of the calculations show the element of a shape that the college students surveyed most preferred what Birkhoff defines as equilibrium and that vertical symmetry does not matter to them. According to the survey conducted, the shape that people prefer the most is  and what people prefer the least is . This survey and other surveys also prove that Birkhoff’s formula has little correlation with real-world observations. With all the corrections to the formulas, it better fits the results of the survey conducted. There are also other factors that contribute to the aesthetic appeal of an object. Finding those factors, on the other hand, is beyond the scope of this paper. But, it could be a potential point of expansion in the field.

A potential test for the results of the paper is to test the new formula against a bigger sample. In the field of artificial intelligence, the sample from the survey is considered the learning sample. Testing the formula calculated using the learning sample on a bigger sample shows how reliable the formula actually is.

Acknowledgments

Professor Harutiun (Harry) Nishanian was my foundation instructor. In this paper, he helped me work through many of the mathematics in the paper. I use up around 2 or 3 hours of his time every week just to talk about the paper. He even consulted his colleagues outside of the college for me.

This paper would not be feasible without the additional help of many of the professors that voluntarily give their time to the paper. Professor Mohsen Janatpour helped during the beginning when I need to narrow my topic down. He was actually the person who recommend that I look into George Birkhoff. Professor Joshua Rhodes helped tremendously in applying statistics in the paper as well as the statistical analysis of the data. He also contributed to the collection of data by conducting surveys across three of his classes, getting 64 responses. Professor Cheyanne Cortez provided an art historian’s perspective on quantifying art and Professor Brooke Candrian provided information about the relationship between art and science in the field of architecture. Professor Jacqueline Yahn and Professor Gary Church each also conducted surveys in one of their own classes.

Bibliography

Birkhoff, G. D. (1933). *Aesthetic measure*. Harvard University Press.

Douchová, Veronika. (2016). *Birkhoff's aesthetic measure*. AUC PHILOSOPHICA ET HISTORICA.

Eysenck, H. J. (1941). *The empirical determination of an aesthetic formula*. Psychological Review, 48(1), 83-92.

Newman, J. R. (2000). Mathematics of Aesthetics. *In The World of Mathematics (Vol. 4)*. Dover Publications.

Schwartz, B. (2004). *The paradox of choice*. Harper Perennial.