

# Does Gender Identification at Birth Correlate with Empathy Level?

**Sophia M. Panigrahi**  
Stanford Online High School  
[skisoph1@ohs.stanford.edu](mailto:skisoph1@ohs.stanford.edu)

## Abstract

The study discussed below is one that centers around the following research question: Is gender identification at birth correlated with empathy level? In understanding the empathy levels of humans, one can gain insight into the topic of empathy with relation to gender and the role that empathy of different genders plays in human society. Moreover, this study holds the possibility of illustrating that transgender people mentally identify with the empathy levels of an alternative gender, contributing to a higher level of understanding from human society. In understanding those who are transgender more fully, studies such as this one could increase societal toleration, allowing transgender people the acceptance they deserve. After collecting dozens of responses via a virtual survey, the collected data was analyzed via the Chi Square Test for Association and the Mann-Whitney U-Test. Ultimately, both tests determined that the collected data was highly statistically significant, indicating that gender identification at birth is correlated with empathy level.

## Key Words

Empathy, gender, statistics, correlation, nature

## 1. Introduction

First, the motivation and basis for this study will be discussed, and some scientific information will be included. Additionally, the hypothesis developed prior to carrying out this study will be detailed.

Often, human females are assumed by society to be the most altruistic or empathetic of the genders (Christov-Moore, Leonardo, *et al.*, 2014). What was intriguing about this commonly believed assumption is that females of specific species in nature tend to be less empathetic than male organisms as shown by the routine consumption of their mates and young (Suttle, 1999). These same behaviors have been predicted and observed in male organisms of specific species more rarely (Revynti, A.M., *et al.*, 2018). Consider the example of the male black widow spider. This illustrates the contrast between the empathy of different genders in the animal kingdom, contributing to the question as to how human empathy appears different between genders. Interested in both the sharp contrast between nature versus human society and the assumption that females are more empathetic, the authors carried out this study in order to determine the answer to the following question: Is gender identification at birth correlated with empathy level? Unlike other studies which have conducted research regarding wild animal behavior with regard to empathy, this study will provide an interesting contrast to studies regarding wild animals. By comparing human empathy to the empathy of wild animals, one can gain insight into the role of empathy and how it differs in human society.

After considering the above question with regard to personal experience, the following hypothesis was developed: Gender identification at birth does correlate with empathy level. A study such as this one is pertinent to society today and is growing more and more relevant to society's discussions. In a world where so many assumptions are simply made without basis, finding the "truth" is a hugely valuable endeavor. In studying this topic, the objective is to determine if the common assumptions made about male, female, or alternate gender empathy levels are accurate, scientifically and statistically. Based on research, my preliminary belief is that the empathy level of males will be lower than that of females.

## 2. Methods

In this portion of my paper, the materials used to carry out this study will be discussed. Further, the method used to carry out this study will be described. This study did not require a great deal of materials to yield statistical data. After collecting statistical data, several procedures were carried, specifically statistical tests.

## 2.1 Materials Used

The author decided that a survey or poll would be the method allowing for optimum data collection as it was a simple method of gathering data in an efficient manner. Knowing that an electronic survey would provide the simplest method for sharing this poll broadly, the survey was generated through a virtual program. By using an electronic poll, the author was able to gather data at a much faster rate, and such data was automatically recorded, simplifying the basic recording process. Ultimately, Google Forms was selected as the platform for constructing this survey, meaning that the only two materials required to carry out this study were a computer and Google Drive account.

## 2.2 Procedure

After some research and material gathering, the process of constructing and circulating this survey, as well as recording the results obtained from the respondents, began. Using a Google Drive account on a desktop computer, an anonymous survey composed of five multiple choice questions was constructed. Of these five questions, two were written in regards to the survey taker's gender at birth and gender at the time that the survey was taken, allowing one to gather critical data regarding. In order to gauge empathy level, the remaining three questions required survey respondents to select one of five possible reactions when presented with certain hypothetical scenarios and corresponding questions. The remaining three questions asked respondents to consider various scenarios that would prompt an empathetic response. The first of these scenarios was "A student who is new to your school walks into the cafeteria alone, looking lonely," and the question prompted respondents to select the option that would most likely describe their reaction. The remaining two scenarios similarly asked respondents to select their most likely reaction. The final two scenarios were "An acquaintance of yours (who goes to school with you) is giving a speech in one of your classes. This person is known for being shy, meaning that this person is nervous when public speaking. When the student speaking makes an error by misreading a word or stuttering, the class bursts into laughter" and "You and several of your friends are watching a tragic movie. At the end of the film, the main character's best friend passes away in a terrible accident, which leaves the main character heartbroken and sobbing."

The first of these scenarios was meant to gauge how respondents would react with regard to empathy to a stranger in a negative situation. The second of these questions aimed to determine how one would react with regard to empathy to an acquaintance in a negative situation. Finally, the third question aimed to gauge how one would react when faced with a negative situation that they are viewing from an objective and unrelated position. By asking questions regarding empathy from various angles, the author hoped to gain an understanding of one's empathy level as a whole and not with regard to just one type of situation.

Each of the five reactions afforded as respondent options corresponded with a specific empathy level on a scale of one to five. The scale of one to five was the selected scale as a number below five increments would not allow for enough variation and a number above five increments would allow for too much variation. By using this range of increments, analyzing statistical data was manageable in that additional increments did not create unnecessary complexity.

Following survey creation, a very simple method for gathering a large sample of survey takers was developed. Using email, and text messaging, the survey link was emailed to a list of personal contacts. Note, it is possible that largely reaching out to personal contacts could possibly lead to bias in that all those surveyed were acquaintances of the author. Sent with the survey link was a brief message including the purpose of the survey and the simple request that every person receiving the message share the survey with additional contacts. Using this method, a total of sixty-seven survey responses were gathered in the time allotted.

Upon closing the survey and completing data collection, the study proceeded with the process of data recording and verification, essential steps in any study. To keep records organized and as part of this scientific methodology, a Google Document comprising a six-column data table was created. This data table illustrated each survey taker's response to the five questions posed on the survey and following these five columns, a new category of 'median empathy level' was constructed with the intent of reducing the responses to a single unweighted empathy 'score.' By using the median empathy level, the author was able to eliminate outliers. The median was selected over the mean because the mean would allow for outliers to be averaged, but by using the median, eliminating outliers was possible to a larger extent. Each 'score' or median empathy level was determined using the suggested empathy levels

indicated by each survey taker’s scenario responses (responses to the last three questions on the survey). In order to record data with accuracy, documentation of all individual results at one time was carried out. Doing so lowered the probability that one would resume documentation at the incorrect place on the list of survey results and thus indicate error in the process of transcription. Further, the data was checked and cross-checked for accuracy. In short, quality data was gathered and recorded using a simple three step process (create, distribute, record) that could be followed if this investigation was ever replicated<sup>1</sup>.

### 3. Data and Statistical Analysis

This section of the paper will describe the data collected as a key step of this study. In order to clearly illustrate this data and its trends, figures are included. Following the recording of all survey data, statistical testing began and will be detailed in this portion of the paper. In order to determine whether or not gender identification at birth is correlated with empathy level, a series of statistical tests (illustrated below) were carried out. As pictured in the table below, basic statistical numbers allow us to determine in a preliminary manner that females, based on the data collected, appear to, on average, have a higher empathy level score. Moreover, the most occurring score on the empathy scale is higher for females than for males who completed the survey.

#### 3.1 Data Presentation

In seeking the answer to the research question, the two variables under observation were as follows: gender identification (at birth) on a nominal scale and empathy level on an ordinal scale. This study aims to discover whether or not empathy level depends on gender identification at birth, and because of this, observed variables cannot be categorized as independent or dependent, which is a minor note to consider.

The data table will not be attached due to the large sample size. The following information provides a summary of the full data set that was collected.

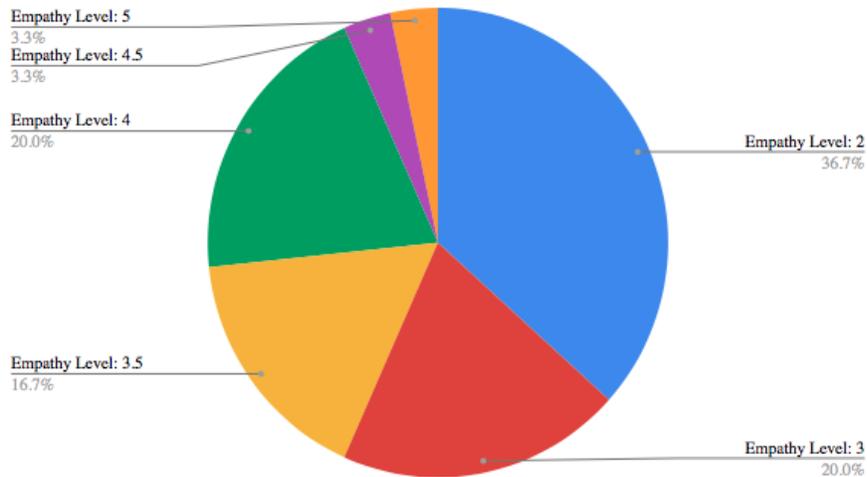
Table 1. Measures of Central Tendency of Collected Data

Measure of Central Tendency	Female	Male
Average Empathy Level	(about) 3.76	(about) 3.03
Standard Deviation	(about) 0.95	(about) 0.92
Mode	3	2

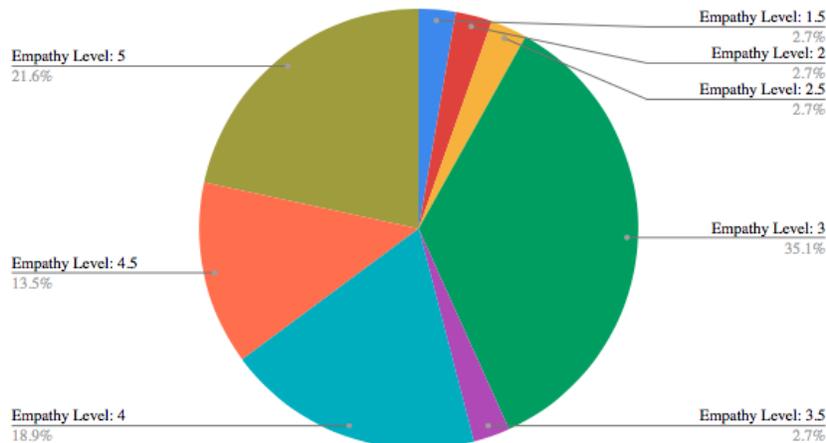
The following charts illustrate the frequency of varying median empathy levels collected from each gender type.

<sup>1</sup> The median empathy levels of some individuals are noted to end in decimals as some respondents opted to respond to only two of the three scenarios listed on the survey (partial completion).

Frequency of Median Empathy Levels in Male Data

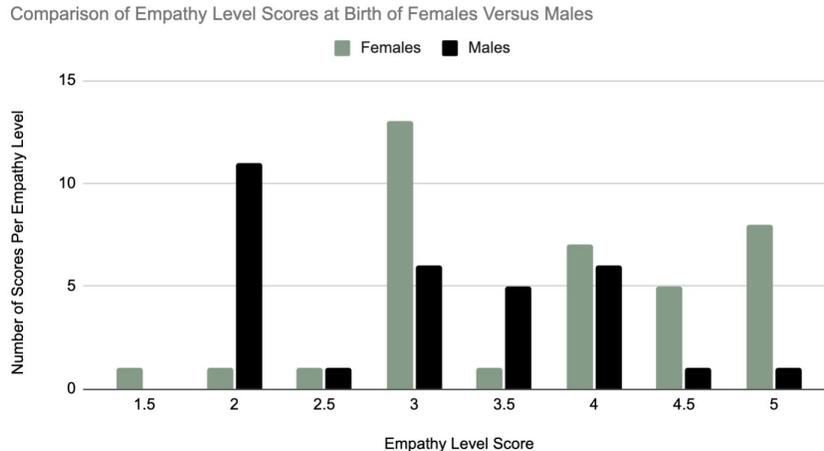


Frequency of Median Empathy Levels in Female Data



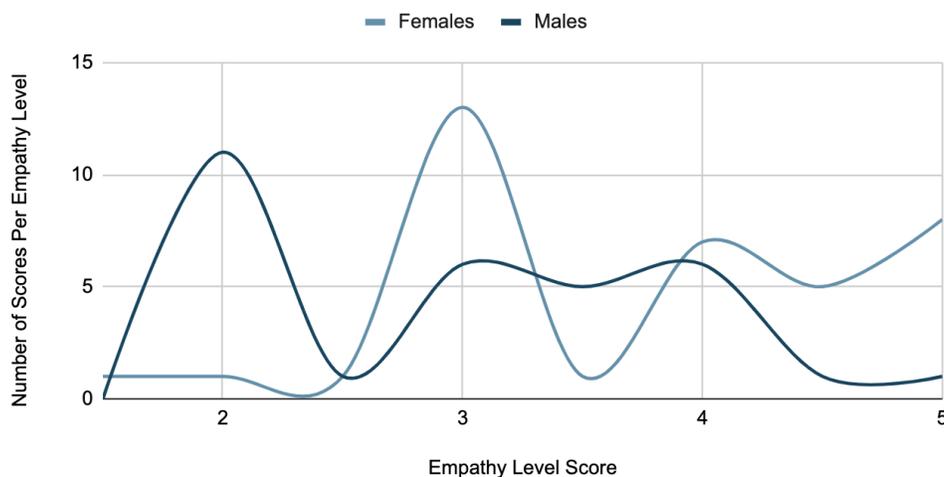
Given the data displayed in the graph above, it is evident that from an initial standpoint, there appears to be some contrast between the median empathy levels of females versus males given the data at hand. As seen in the graph of male data, 0% of male survey takers scored a five on the empathy scale, the highest possible score, but as seen in the graph of female data, 21.6% of females scored a five on the empathy scale. Given this information alone, there appears to be some contrast between the data corresponding to each group. When considering additional data, as seen in both graphs, a larger percentage of males, in comparison to females, scored a three of two on the empathy scale. Based on initial observation, differences between the data from each group appears evident in that females appear to score higher on the empathy scale than males based on preliminary observation. Additionally, the data displayed in Table 1 highlights the higher mean (3.76) and higher mode (3) of female data in comparison to the lower mean (3.3) and lower mode (2) of male data.

While the complete data table<sup>2</sup> (not pictured) does illustrate the gender identification of survey takers at the time of survey completion, this information was found to have little relevance as it was noted that there was no significant change between the collective gender identifications of survey takers at birth versus the gender identifications at the time of survey completion. This data, though not pictured, helps to represent all aspects of survey results and shows that the genders of respondents for this particular study seem to be somewhat ‘fixed’ given the results of this survey. Though this study was not structured to answer this question, the results themselves appear to suggest an interesting pattern that could form the basis for subsequent studies in the future.



The bar graph above additionally displays the data gathered through the use of the online survey. Through this method of graphical representation, the skew of female empathy level scores toward the right, or toward higher empathy level scores, is increasingly evident. Moreover, one can visually understand the skew of male empathy level scores toward the left, or toward lower empathy level scores. This indicates that females tend to score higher on the empathy level scale than males, given the data at hand.

Comparison of Empathy Level Scores at Birth of Females Versus Males



<sup>2</sup> Please contact author to view complete data table

The method of graphical representation in the above graph illustrates the fluctuation of empathy level scores with regard to gender. Following the female data curve, one can determine that in general, females score very rarely below an empathy level score of three but appear to score much more commonly at an empathy level score of three or above. Following the male data curve, one can determine that in general, males score very commonly below an empathy level score of three but appear to score much less commonly at an empathy level score of three or above. This graphical representation using curves illustrates the opposite tendencies of each gender with regard to empathy level scores. The female data set appears to reflect a higher empathy level score than the male data set.

### 3.2. Chi Square Test for Association

In conducting this test, one will determine if there is an association/correlation between the two variables under observation (in this case gender identification at birth and empathy level), which, in fact, directly addresses the research interest. The Chi Square Test for Association is the appropriate application in this case for several reasons. Firstly, this particular test calls for two variables in which each variable is categorized into two or more groups, which is met by the data collected. Additionally, because this test is based more on the frequency of data rather than the data itself, it is possible for the Chi Square Test for Association to handle all four measurement scales (including nominal and ordinal data). The null hypothesis for this experiment is as follows: there is no association between gender identification at birth and empathy level. The alternative hypothesis for this experiment is as follows: there is an association between gender identification at birth and empathy level.

### 3.3 Trial #1

Table 2. Data Used for Trial #1 of the Chi Square Test for Association

Empathy Level	Female	Male
0.5-1	0	0
1.5-2	2	11
2.5-3	14	6
3.5-4	8	11
4.5-5	13	2

The resulting Chi-Square value of 17.43, as seen in Table 2, exceeded both the 95% (9.49) and 99% (13.28) significance thresholds, meaning that the correlation between gender identification at birth and empathy level is statistically highly significant, and the alternative hypothesis should be accepted<sup>3</sup>. This test was valuable in that it allowed the author to draw the preliminary conclusion that there was a statistically highly significant difference between the empathy levels of males and females. In gaining this result, one is able to gain insight into the role of empathy level with regard to gender identification in society. Moreover, results such as these can contribute greatly to discussions aimed towards ending the discrimination of those who are persecuted due to gender identification.

### 3.4 Trial #2

Because certain median empathy levels ended in decimals, it was necessary to create empathy level intervals, which is the most appropriate method to handle the associated data. In order to ensure that these intervals did not alter the results of the statistical tests, intervals that ‘round’ decimal values up/down were tested. By using round decimal values, the author was able to use interval scale data, simplifying the process of carrying out relevant statistical tests. In other words, one set of intervals begins with a value of 0.5, while the other begins with 1.0 (ie: “rounding” up/down). There are no intermediate decimal values.

Table 3. Data Used for Trial #2 of the Chi Square Test for Association

<sup>3</sup> Please note that the Degrees of Freedom for both trials of the Chi Square Test for Association is 4.

Empathy Level	Female	Male
1-1.5	1	0
2-2.5	2	11
3-3.5	14	11
4-4.5	12	7
5-5.5	8	1

The resulting Chi-Square value of 13.77 exceeded both the 95% (9.49) and 99% (13.28) significance thresholds, meaning that that the correlation between gender identification at birth and empathy level is statistically highly significant, and again, that the alternative hypothesis should be accepted. The result of this test is particularly important as it proves that the formation of empathy level intervals did not alter the results of the statistical tests, suggesting stronger validity of the observations. Moreover, this test highlights the reliability of the data at hand.

### 3.5 Mann-Whitney U-Test

By conducting the Mann-Whitney U-Test, it will be determined if there is a significant difference between the average<sup>4</sup> empathy levels of males versus females. From here, one would be able to make further conclusions based on the results of this test. Because the data collected falls on the ordinal and nominal scales, it is only logical to use the Mann-Whitney U-test as it is a nonparametric test. The null hypothesis for this experiment is as follows: there is no difference between the average empathy levels of males versus females. The alternative hypothesis for this experiment is as follows: there is a significant difference between the average empathy levels of males versus females.

Table 4. Data Used for the conduction of the Mann-Whitney U-Test

Empathy Level	Female	Male
1	0	0
1.5	1	0
2	1	11
2.5	1	0
3	13	6
3.5	1	5
4	7	6
4.5	5	1
5	8	1

<sup>4</sup> When the word “average” is used in this portion of the paper, it is meant to convey “the average of the empathy level medians” as opposed to an average of means, which is not possible to assess accurately on an ordinal scale.

Due to the fact that both sample sizes of the data were greater than or equal to thirty<sup>5</sup>, it was possible to analyze the significance level of the data using z-score values. It was decided to analyze z-scores rather than critical values as the sample sizes were so large that upon searching, the appropriate table illustrating the correct critical value could not be found. The resulting z-score value of 2.72 exceeded both the 95% (1.96) and 99% (2.56) significance thresholds, meaning that there is a statistically highly significant difference between the average empathy of females versus males. This is important because the statistically highly significant difference between the empathy of females versus males highlights the difference in empathy level between the two groups. In understanding this, one is able to continue with further studies comparing human empathy to animalistic empathy.

#### 4. Conclusion

Having completed this study, the statistical findings of this study in relation to the central research question<sup>6</sup> will now be discussed. As stated above, results of the Chi-Square Test for Association were statistically highly significant regardless of varying data intervals. The results demonstrate a statistically highly significant correlation between gender identification at birth and empathy level supporting my original hypothesis. Slightly more abstract, the Mann-Whitney U-Test result indicated a statistically highly significant difference between the average empathy levels of males versus females. By applying the result of this Mann-Whitney U-Test to a previous observation that females score higher empathy levels than males, it can be concluded that females generally retain a higher empathy level than males.

If this study were to be replicated, two elements of the study's design would likely be altered. Given that the objective of this study was to determine if birth gender identification is correlated with empathy level, it would have been more valuable to have collected data outside of identification as male or female alone. Specifically, it is known that gender identification is likely not a simple 'black or white' data point and, as is the case with other human characteristics, gender identification may be better conceptualized as 'multiple shades of gray.' Empirically testing some of these alternative 'shades' would have proved invaluable. A wider respondent pool better reflecting the beauty and variability of the human condition, would have provided better insight into gender identifications with empathy level. In subsequent studies, though requiring more resources and funds, it would be beneficial to broaden the study's sample as a means of encompassing more gender identification subtypes. In order to do this, publicizing this survey on various internet sites and media outlets would be necessary.

Additionally, while decimal medians did not interfere with the integrity of the statistical tests, it would be beneficial to avoid decimals in order to steer clear of complications. Simply altering the original survey to require a response to all three scenarios described could have an impact in that each survey taker's median would be entirely clear.

#### 5. Future Research

As this study was carried out, a second research question developed: Does a change in gender correlate with a change in empathy level? In the future, a study expanding on this topic could be carried out. Following a similar process, one would issue a survey, recording its results. One would publicize this survey more widely in order to gather data that better reflects a change in gender. For example, does a respondent who undergoes sexual reassignment surgery maintain their prior empathy scores or do scores change? If there is change, how do variables such as age of the transition, psychological support, etc. influence the results?

In society today, assumptions based on gender, more often than not, are baseless in that they have not been scientifically proven (Ahnesjö, I., *et al.*, 2020). Studies such as this one investigate the root of so many assumptions that are made solely based on gender identification (Ahnesjö, I., *et al.*, 2020). Moreover, human society is made up of not only those who identify with their birth gender, but also of transgender humans, those who identify with one or more alternative gender(s). Transgender people have been baselessly criticized for their personal gender identification decisions. Beyond this criticism, transgender people have also been systematically persecuted by unjust laws that restrict the rights of those who are transgender (Divan, Vivek, *et al.*, 2016).

<sup>5</sup> The female sample data size was 37. The male sample data size was 30. The U/UA for this test was 338.5.

<sup>6</sup> Is gender identification at birth correlated with empathy level?

In short, this study provided the opportunity to apply biological and statistical concepts to a realistic and pertinent topic with real world implications for society today. Future research building on the study discussed throughout this paper could contribute to the end of the baseless persecution of transgender people. Studies such as the one described in the above subsection could possibly demonstrate that those who are transgender mentally identify with the empathy levels of an alternative gender. By studying the relationship between changes in gender and empathy level, human society would be able to better understand those who are transgender. The increased understanding that comes from such scientific studies could likely lead to the increased level of tolerance from human society that transgender people deserve.

## 6. Acknowledgements

The author of this paper would like to thank VassarStats for the use of the organization's Chi-Square Test for Association Calculator and Mann-Whitney U-Test Calculator. The measures of central tendency of the collected data were also calculated via VassarStats. The author of this paper would also like to thank Jim Fowler, Lou Cohen, and Phil Jarvis for the use of their book, *Practical Statistics For Field Biology*. *Practical Statistics For Field Biology* provided the critical value reference tables necessary for analyzing the collected data via the Chi-Square Test for Association and the Mann-Whitney U-Test.

## References

- Ahnesjö, I., Brealey, J.C., Günter, K.P. *et al.* Considering Gender-Biased Assumptions in Evolutionary Biology. *Evol Biol* 47, 1–5 (2020).  
<https://doi.org/10.1007/s11692-020-09492-z>
- Christov-Moore, Leonardo et al. “Empathy: gender effects in brain and behavior.” *Neuroscience and biobehavioral reviews* vol. 46 Pt 4, Pt 4 (2014): 604-27.  
doi:10.1016/j.neubiorev.2014.09.001
- Divan, Vivek et al. “Transgender social inclusion and equality: a pivotal path to development.” *Journal of the International AIDS Society* vol. 19,3 Suppl 2 20803. 17 Jul. 2016, doi:10.7448/IAS.19.3.20803
- Revynti, A M et al. “Gender-specific differences in cannibalism between a laboratory strain and a field strain of a predatory mite.” *Experimental & applied acarology* vol. 74,3 (2018): 239-247. doi:10.1007/s10493-018-0232-4
- Suttle, Kenwyn B. *The Evolution of Sexual Cannibalism*. [ib.berkeley.edu/courses/ib160/past\\_papers/suttle.html](http://ib.berkeley.edu/courses/ib160/past_papers/suttle.html).

## Biography

**Sophia M. Panigrahi** is a student at Stanford Online High School. In addition to her passion for biological and medical studies, Sophia is also profoundly interested in a variety of philosophical topics. Currently, Sophia hopes to publish a work-in-progress paper that centers around the philosophical nature of time, and she recently published a paper to the *Journal of Double Star Observation*. Working with her two sisters, Sophia is also pursuing her interests in art and website design as she aims to launch a charity-based website. Further, Sophia has opened a small, art-based business in recent months. When not focused on academic topics or passion projects, Sophia is dedicated to competitive fencing. Sophia would like to thank any and all readers for their interest in her paper!