A black and white circular pattern

Description automatically generated with medium confidence

**Revisiting McManus’ 1985 Handedness Model**

**Tomer Oron**

322549312

**Supervisor:**

Dr. Yoav Ram School of Zoology

Faculty of Life Sciences Tel Aviv University

December 2023

Contents

[Introduction 3](#_Toc153989599)

[Model 4](#_Toc153989600)

[Data 6](#_Toc153989601)

[Familial data 6](#_Toc153989602)

[Twin data 11](#_Toc153989603)

[Correction to the prediction given the datasets 13](#_Toc153989604)

[Model fitting 13](#_Toc153989605)

[Statistical analysis 15](#_Toc153989606)

[Results 15](#_Toc153989607)

[Discussion 19](#_Toc153989608)

[APPENDIX 1: CORRECTION MATRICES 22](#_Toc153989609)

[Correcting triplets datasets 22](#_Toc153989610)

[Correcting families with multiple children 24](#_Toc153989611)

[Correcting for twins 24](#_Toc153989612)

[APPENDIX 2: PROBABILITIES AND LIKELIHOOD FUNCTIONS 25](#_Toc153989613)

[Likelihood functions 28](#_Toc153989614)

[Bibliography 37](#_Toc153989618)

**Table of tables**

[Table 1. The expected probabilities of presenting right- and left- handedness by genotype 3](#_Toc153726527)

[Table 2. Mating data from Table 2 in McManus (1985). 5](#_Toc153726528)

[Table 3. Mating data from Table 3 in McManus (1985). 7](#_Toc153726529)

[Table 4. Observations of twins taken from Table 5 in McManus (1985). 8](#_Toc153726530)

[Table 5. Comparison of maximum log-likelihood of the present model, and the model of Mcmanus 11](#_Toc153726531)

**Table of equations**

[Equation 1. The frequency of allele C in the population 18](#_Toc145611532)

[Equation 2. probabilities of progeny expressing left-handedness given parents handedness 18](#_Toc145611533)

[Equation 3. probability function of twins expressing phenotypes 19](#_Toc145611534)

[Equation 4. likelihood function for triplets 20](#_Toc145611535)

[Equation 5. likelihood function for families with multiple offspring 20](#_Toc145611536)

[Equation 6. likelihood function for twins 20](#_Toc145611537)

Abstract

This project presents a replication and critical evaluation of a previously published genetic model for human handedness by McManus (1985). McManus' model postulates a single gene with two alleles as the determinant of handedness and serves as a foundational framework for understanding the genetic foundations of this phenomenon who serves? The gene or the model?. Through our replication efforts, we successfully estimated parameters closely mirroring McManus' original findings, while also rectifying reporting errors in the original statistical values, enhancing the model accessibility and applicability. Our revised model met McManus' established criteria for a genetic model of handedness, fostering continued exploration into the genetic foundations of handedness while reducing complexities and barriers in its utilization.

## Introduction

Handedness, most commonly defined as the preferred hand used in one-handed tasks [(Porac 2016)](https://sciwheel.com/work/citation?ids=15365161&pre=&suf=&sa=0&dbf=0), has intrigued researchers for decades [(McManus 2019)](https://sciwheel.com/work/citation?ids=14962387&pre=&suf=&sa=0&dbf=0). The majority of the human population exhibits right-handedness, a phenomenon evident not only in modern society but also in artistic representations spanning millennia [(Coren and Porac 1977)](https://sciwheel.com/work/citation?ids=10047819&pre=&suf=&sa=0&dbf=0). Moreover, archaeological discoveries provide evidence that this bias towards right-handedness has been a longstanding trait within human societies, with indications dating back to Neanderthals [(Conard 2011)](https://sciwheel.com/work/citation?ids=15365201&pre=&suf=&sa=0&dbf=0). This consistent occurrence of right-handedness across various historical periods and even in our ancient ancestors strongly hints at the genetic keystones of handedness, reinforcing the hypothesis that genetics play a significant role in determining an individual's preferred hand.

In the 1970s, Marian Annet introduced the "right shift theory" [(Annett 1972)](https://sciwheel.com/work/citation?ids=6366514&pre=&suf=&sa=0&dbf=0) as a novel genetic model for understanding handedness. Unlike earlier models suggesting a single gene with two alleles dictating handedness, with the recessive allele corresponding to left-handedness, Annet's model incorporated an element of chance into handedness determination. Her model proposed the existence of a genetic mechanism that shifts the distribution of handedness within the human population toward the right hand. Individuals inheriting the gene for right-shift are likely to present right-handedness, while for those without this gene, handedness is determined by chance, with an equal probability for both right- and left-handedness.

Building on Annet's framework, McManus proposed a similar genetic model [(McManus 1985)](https://sciwheel.com/work/citation?ids=15020506&pre=&suf=&sa=0&dbf=0), although with slight variations: McManus' model also used the element of chance in handedness determination but suggested that genes directly control hand preference rather than influencing brain lateralization, as Annet had hypothesized. To this day, researchers continue to engage in debates over the precise mechanisms governing handedness determination, and McManus' model remains one of the leading theories in this field.

Here, we revisit McManus' genetic model of handedness using present-day computational tools to reaffirm the model's validity and critically examine its underlying parameters. Initially, we will describe the specifics of McManus' model, **explaining its requirements and parameters**. Secondly, we will examine the **datasets** used to study the model, as well as the **correction method** applied to ensure the integrity of the analysis. Thirdly, we will describe **the approach to estimating** the model parameters through maximum likelihood estimation and **test the goodness of fit** of the model to the available data. Finally, we will **present the results of our replication**.

## Model

In 1985, McManus introduced a genetic model of handedness, suggesting that a single gene with two alleles, D and C, controls handedness. When allele D is homozygous, it leads to a 100% occurrence of right-handed individuals, while allele C, when homozygous, results in a 50% chance of either handedness. The primary goal of this model is to calculate the expected proportion of left-handers among heterozygotes, denoted as p(L│DC), and determine the frequency of allele C within the population. Table 1 outlines the probabilities of presenting left-handedness and right-handedness for the different genotypes.

Table 1. The expected probabilities of presenting right- and left- handedness by genotype

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

The model operates under several assumptions. First, there is no assortative mating, indicating that individuals do not choose mates based on the phenotype of handedness. Second, genes control the phenotypes of twins similarly to the way they control the phenotypes of singletons. Third, each individual is either left-handed or right-handed, thus resulting in bimodal distribution. Fourth, the cause to difference in the observed left-handedness rates across generations and studies is due to variation in the criteria for determining left-handedness by the researchers collecting the data. Fifth, the model assumes uniform allele frequencies throughout all studies.

McManus required the model to answer a number of requirements:

1. Familial Pattern in Handedness: The model must account for the observed familial pattern in handedness. This pattern indicates that left-handers make up approximately 10% of the children of two right-handed parents, 20-25% of the children of one right-and one left-handed parent, and approximately 40% of the children of two left-handed parents.
2. Twins' discordance rate**:** The model must account for the literature that indicates a high proportion of monozygotic twin pairs that display discordant handedness, with one being left-handed and the other right-handed.
3. Different incidence of handedness: The model must incorporate the observed differences in the prevalence of left-handedness among different populations and generations.
4. Compatibility with other biological asymmetries: The model should align with the known inheritance patterns of other biological asymmetries, such as *situs inversus* (the inverted position of chest and abdominal organs), hand clasping*,* and arm folding.
5. Biologically convincing: A genetic model can be applied to any dataset, provided that the dataset includes a diverse set of genetic variations (alleles) at different genetic loci and considers the varying penetrance of these alleles on the trait.

The model uses two parameters to describe the population:

1. : The true proportion of left handers in the population. As indicated by the third requirement of the model, the *observed* proportion of left handers varies across studies and generations. Therefore, it is necessary to determine the true proportion of left-handers in order to accurately account for this variability. It is assumed that the true proportion of left-handers in the population is constant across all populations studied in this research.
2. : The probability of left-handed phenotype in heterozygotes. This parameter is essential for characterizing the allele frequencies within the population and understanding the genetic system that governs handedness.

## Data

McManus used two types of datasets: one describing familial data, including parental phenotypes and the phenotypes of the progeny resulting from such matings; and another presenting twin data. The twin dataset displayed observed numbers of twin pairs, both monozygotic (MZ) and dizygotic (DZ) presenting the possible phenotypes of handedness (L×L, R×L, R×R).

### Familial data

The familial data is presented in two tables in McManus (1985): the first table presents triplets (single offspring and parents) and the second table presents families with multiple children (siblings and parents). The first table summarizes data of measured handedness from twelve distinct studies, that describe the incidence of left-handedness in the progeny resulting from different matings (L×L, R×L, R×R). These studies contained families, demonstrating an average left-handedness of in progeny and left-handedness in parents. Specifically, the average left-handedness measured in offspring of two right-handed parents (R×R) is , while it was in offspring of one right-handed and one left-handed parents (R×L) and for offspring of two left-handed parents (L×L). These statistics align with the known data regarding handedness, as stated in the first requirement of the model. Table 2 contains the observed number of offspring with each phenotype (right- or left-handed) given the phenotypes of the parents, as described in McManus (1985).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| study |  |  |  | |  | |  | | (df=1) |
| R | L | R | L | R | L |
| Ramaley (1913) | 0.1556 | 0.0809 | 841 | 115 | 113 | 54 | 1 | 7 | 18.897\*\*\* |
| *822.76* | *133.24* | *126.96* | *40.04* | *5.28* | *2.72* |
| Chamberlain (1928) | 0.0477 | 0.0333 | 6917 | 308 | 411 | 53 | 18 | 7 | 1.709 |
| *6912.7* | *312.3* | *412.74* | *51.26* | *20.56* | *4.44* |
| Rife (1940) | 0.0877 | 0.0450 | 1842 | 151 | 140 | 34 | 5 | 6 | 3.142 |
| *1837.84* | *155.16* | *141.42* | *32.58* | *7.74* | *3.26* |
| Merrell (1975) | 0.2363 | 0.1540 | 140 | 34 | 33 | 20 | 8 | 2 | 4.333\* |
| *135.44* | *38.56* | *38.73* | *14.27* | *6.83* | *3.17* |
| Annett (1973) | 0.1063 | 0.0407 | 6206 | 669 | 471 | 125 | 5 | 1 | 0.77 |
| *6203.64* | *671.36* | *474.23* | *121.77* | *4.13* | *1.87* |
| Ferronato et al. (1974) | 0.0976 | 0.0976 | 154 | 11 | 31 | 9 | 0 | 0 | 1.253 |
| *151.77* | *13.23* | *33.23* | *6.77* |
| Mascie-Taylor (unpub) | 0.0831 | 0.0930 | 232 | 17 | 41 | 7 | 3 | 1 | 0.109 |
| *232.71* | *16.29* | *40.31* | *7.69* | *2.98* | *1.02* |
| Chaurasia & Goswami (unpub) | 0.1407 | 0.0660 | 1060 | 144 | 122 | 46 | 3 | 4 | 3.745 |
| *1051.36* | *152.64* | *129.00* | *39.00* | *4.64* | *2.36* |
| Annett (1978) | 0.0850 | 0.0545 | 1656 | 130 | 170 | 40 | 4 | 0 | 2.855 |
| *1655.76* | *130.24* | *171.42* | *38.58* | *2.82* | *1.18* |
| Carter-Saltzmann (1980) | 0.1300 | 0.0750 | 303 | 37 | 45 | 15 | 0 | 0 | 0.36 |
| *301.29* | *38.71* | *46.71* | *13.29* |
| Coren & Porac (1980) | 0.1842 | 0.0800 | 315 | 68 | 57 | 16 | 0 | 0 | 1.093 |
| *318.46* | *64.54* | *53.54* | *19.46* |
| McGee & Cozad (1980) | 0.2415 | 0.1825 | 848 | 211 | 325 | 150 | 30 | 22 | 13.487\*\*\* |
| *318.46* | *64.54* | *53.54* | *19.46* | *318.46* | *64.54* |
| The numbers in italic are expected values for model with . | | | | | | | | | |

Table 2. The results of 12 studies on the incidence of right- and left-handed offspring of R×R, R×L and L×L parents. Mating data from Table 2 in McManus (1985).

The second table in McManus (1985) presents the results of two surveys conducted by McManus among students at the University of Cambridge. Students were asked to provide information regarding their own handedness, their siblings' handedness, as well as their parents' and grandparents' handedness. The first survey (referred as ICM1) was conducted in May 1977 among undergraduate students. Handedness was classified based on the hand used for writing, except for those who wrote using their right hand due to being forced to switch from writing using their left hand. Such cases were classified as left-handers. The second survey (referred as ICM2) took place in June 1977 among graduate students on the eve of the graduation ceremony. The classification of handedness in this questionnaire was similar to the first survey. From these questionnaires, three datasets were obtained: one for students and their siblings, considering their parents' handedness, and two datasets for the parents' and their siblings' handedness given the grandparents' handedness, with distinction between paternal and maternal family. The data from both surveys are presented in Table 3.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Family size |  | | | | | |  | | | | | |  | | | | | |  |
| 0 | 1 | 2 | 3 | 4 | 5 | 0 | 1 | 2 | 3 | 4 | 5 | 0 | 1 | 2 | 3 | 4 | 5 |
| ICM1 | 0.1518 | 0.1005 | 1 | 58 | 9 |  |  |  |  | 14 | 5 |  |  |  |  | 0 | 0 |  |  |  |  | 40.161  df = 44 |
| *57.92* | *9.08* |  |  |  |  | *14.88* | *4.12* |  |  |  |  |  |  |  |  |  |  |
| 2 | 211 | 57 | 3 |  |  |  | 35 | 16 | 5 |  |  |  | 0 | 0 | 1 |  |  |  |
| *204.02* | *60.52* | *6.47* |  |  |  | *34.73* | *18.27* | *3.00* |  |  |  | *0.50* | *0.40* | *0.10* |  |  |  |
| 3 | 123 | 63 | 6 | 0 |  |  | 22 | 24 | 6 | 1 |  |  | 2 | 1 | 1 | 0 |  |  |
| *126.69* | *53.56* | *10.75* | *1.00* |  |  | *26.29* | *19.75* | *6.18* | *0.78* |  |  | *1.45* | *1.66* | *0.76* | *0.13* |  |  |
| 4 | 70 | 39 | 6 | 1 | 0 |  | 8 | 8 | 7 | 1 | 0 |  | 0 | 0 | 1 | 0 | 0 |  |
| *67.47* | *36.29* | *10.29* | *1.80* | *0.15* |  | *9.61* | *9.18* | *4.12* | *0.99* | *0.11* |  | *0.27* | *0.38* | *0.25* | *0.08* | *0.01* |  |
| 5 | 9 | 12 | 6 | 2 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *14.94* | *9.62* | *3.44* | *0.85* | *0.14* | *0.01* | *0.65* | *0.74* | *0.43* | *0.15* | *0.03* | *0.00* | *0.20* | *0.34* | *0.28* | *0.14* | *0.04* | *0* |
| ICM2 (propositi) | 0.134 | 0.101 | 1 | 134 | 15 |  |  |  |  | 17 | 9 |  |  |  |  | 1 | 0 |  |  |  |  | 37.665  df = 27 |
| *131.52* | *17.48* |  |  |  |  | *20.81* | *5.19* |  |  |  |  | *0.72* | *0.28* |  |  |  |  |
| 2 | 91 | 22 | 2 |  |  |  | 19 | 3 | 3 |  |  |  | 0 | 0 | 0 |  |  |  |
| *90.26* | *22.50* | *2.24* |  |  |  | *16.18* | *7.64* | *1.17* |  |  |  |  |  |  |
| 3 | 22 | 7 | 0 | 2 |  |  | 6 | 11 | 3 | 0 |  |  | 0 | 0 | 0 | 0 |  |  |
| *21.77* | *7.67* | *1.43* | *0.13* |  |  | *10.58* | *7.10* | *2.07* | *0.25* |  |  |  |  |
| 4 | 10 | 5 | 0 | 0 | 0 |  | 3 | 1 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 |  |
| *9.48* | *4.21* | *1.10* | *0.19* | *0.02* |  | *1.74* | *1.48* | *0.62* | *0.14* | *0.01* |  | *0.29* | *0.39* | *0.24* | *0.08* | *0.01* |  |
| 5 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *2.29* | *1.21* | *0.40* | *0.09* | *0.01* | *0* |
| ICM2 maternal | 0.0893 | 0.0609 | 1 | 74 | 4 |  |  |  |  | 6 | 2 |  |  |  |  | 0 | 0 |  |  |  |  | 47.579\*  df = 28 |
| *72.09* | *5.91* |  |  |  |  | *6.51* | *1.49* |  |  |  |  |  |  |  |  |
| 2 | 107 | 18 | 3 |  |  |  | 11 | 3 | 0 |  |  |  | 2 | 0 | 0 |  |  |  |
| *110.20* | *16.21* | *1.60* |  |  |  | *9.33* | *4.10* | *0.56* |  |  |  | *0.99* | *0.82* | *0.19* |  |  |  |
| 3 | 81 | 16 | 4 | 0 |  |  | 16 | 1 | 0 | 2 |  |  | 0 | 0 | 0 | 0 |  |  |
| *81.56* | *16.18* | *3.00* | *0.26* |  |  | *10.46* | *6.63* | *1.72* | *0.19* |  |  |  |  |
| 4 | 31 | 10 | 1 | 0 | 0 |  | 0 | 0 | 0 | 2 | 0 |  | 0 | 0 | 0 | 0 | 0 |  |
| *32.01* | *7.64* | *2.00* | *0.33* | *0.03* |  | *0.91* | *0.75* | *0.28* | *0.06* | *0.01* |  |  |
| 5 | 19 | 7 | 1 | 1 | 0 | 0 | 3 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *20.24* | *5.46* | *1.80* | *0.42* | *0.06* | *0.00* | *1.91* | *1.89* | *0.89* | *0.26* | *0.05* | *0.00* |
| ICM2 paternal | 0.091 | 0.047 | 1 | 86 | 4 |  |  |  |  | 8 | 1 |  |  |  |  | 0 | 0 |  |  |  |  | 28.501 df = 34 |
| *82.77* | *7.23* |  |  |  |  | *7.29* | *1.71* |  |  |  |  |  |  |  |  |
| 2 | 100 | 27 | 0 |  |  |  | 7 | 5 | 0 |  |  |  | 0 | 1 | 0 |  |  |  |
| *108.31* | *16.97* | *1.71* |  |  |  | *7.95* | *3.55* | *0.50* |  |  |  | *0.50* | *0.41* | *0.09* |  |  |  |
| 3 | 65 | 11 | 2 | 0 |  |  | 4 | 3 | 2 | 0 |  |  | 0 | 1 | 0 | 0 |  |  |
| *62.14* | *13.15* | *2.49* | *0.22* |  |  | *4.91* | *3.16* | *0.84* | *0.09* |  |  | *0.35* | *0.43* | *0.19* | *0.03* |  |  |
| 4 | 39 | 11 | 2 | 0 | 0 |  | 2 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 |  |
| *38.95* | *9.92* | *2.65* | *0.45* | *0.04* |  | *1.35* | *1.12* | *0.42* | *0.09* | *0.01* |  |
| 5 | 13 | 9 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *16.29* | *4.70* | *1.58* | *0.38* | *0.06* | *0.00* | *0.38* | *0.38* | *0.18* | *0.05* | *0.01* | *0* |
| The numbers in italic are expected values for model with . | | | | | | | | | | | | | | | | | | | | | | |

Table 3. The numbers of families with particular incidence of left-handed offspring by family size and parental handedness divided to 4 data sets. Mating data from Table 3 in McManus(1985).

### Twin data

The data on twin phenotypes in McManus (1985) paper was presented in a single table, distinguishing between MZ and DZ twins and showing the observed number of twins matching each phenotype, without considering parental phenotypes. These data were collected from thirteen different twin studies, containing a total of pairs of MZ twins with a left-handedness rate of and a discordance rate of . Additionally, the data include pairs of DZ twins with a left-handedness rate of and a discordance rate of . These data are consistent with the previous knowledge about handedness as described in the third requirement of the model. Observed data from these studies are presented in Table 4.

Table 4. The observed numbers of R-R, R-L, and L-L pairs of MZ and DZ twins from 13 differant studies. Observations taken from Table 5 in McManus (1985).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| study | MZ twins | | | | | DZ twins | | | | |
|  |  |  |  | df=1 |  |  |  |  | df=1 |
| Wilson & Jones (1932) | 0.1071 | 56 | 13 | 1 | 0.769 | 0.1138 | 97 | 24 | 2 | 0.166 |
| *56.88* | *11.24* | *1.88* | *97.52* | *22.95* | *2.52* |
| Stocks (1933) | 0.0952 | 35 | 6 | 1 | 0.003 | 0.1064 | 76 | 16 | 2 | 0.037 |
| *35.04* | *5.92* | *1.04* | *75.79* | *16.43* | *1.79* |
| Newman et al. (1937) | 0.19 | 34 | 13 | 3 | 0.223 | 0.1100 | 39 | 11 | 0 | 2.399 |
| *33.44* | *14.13* | *2.44* | *39.99* | *9.03* | *0.99* |
| Bouterwek (1938) | 0.1885 | 80 | 38 | 4 | 1.122 | 0.1714 | 23 | 12 | 0 | 3.203 |
| *81.88* | *34.24* | *5.88* | *24.26* | *9.48* | *1.26* |
| Rife (1940) | 0.1188 | 176 | 41 | 6 | 0.061 | 0.1541 | 104 | 39 | 3 | 0.804 |
| *176.48* | *40.04* | *6.48* | *105.47* | *36.06* | *4.47* |
| Thyss (1946) | 0.1845 | 72 | 24 | 7 | 1.634 | 0.1628 | 60 | 24 | 2 | 0.43 |
| *69.82* | *28.35* | *4.82* | *60.86* | *22.28* | *2.86* |
| Rife (1950) | 0.1283 | 261 | 76 | 6 | 3.749 | 0.1161 | 164 | 45 | 2 | 2.281 |
| *265.66* | *66.68* | *10.66* | *166.43* | *40.14* | *4.43* |
| Dechaume (1957) | 0.2424 | 19 | 12 | 2 | 0.079 | 0.1970 | 21 | 11 | 1 | 0.281 |
| *19.30* | *11.39* | *2.30* | *21.49* | *10.03* | *1.49* |
| Zazzo (1960) | 0.1332 | 199 | 51 | 9 | 0.087 | 0.1090 | 264 | 69 | 2 | 5.72\*\* |
| *198.34* | *52.32* | *8.34* | *268.54* | *59.93* | *6.54* |
| Carter-Saltzmann et al. (1976) | 0.1711 | 132 | 46 | 9 | 0.235 | 0.1932 | 115 | 54 | 7 | 0.098 |
| *130.95* | *48.10* | *7.95* | *115.67* | *52.66* | *7.67* |
| Loehlin & Nichols (1976) | 0.1411 | 380 | 123 | 11 | 4.402\* | 0.1111 | 261 | 70 | 2 | 5.917\*\* |
| *386.52* | *109.96* | *17.52* | *265.64* | *60.72* | *6.64* |
| Springer & Searleman (1978) | 0.1667 | 53 | 19 | 3 | 0.004 | 0.1596 | 35 | 9 | 3 | 1.997 |
| *53.09* | *18.83* | *3.09* | *33.52* | *11.97* | *1.52* |
| NCDS (unpublished) | 0.1512 | 32 | 9 | 2 | 0.184 | 0.1477 | 66 | 18 | 4 | 1.182 |
| *31.58* | *9.84* | *1.58* | *64.54* | *20.93* | *2.54* |
| The numbers in italic are expected values for model with . | | | | | | | | | | |

## Correction to the prediction given the datasets

Evident from the tables is the variation in left-handedness rates across different studies and generations. McManus suggested that disparities in study outcomes might have been due to different criteria used to measure handedness. Therefore, when making predictions based on model parameters and the frequency of left-handedness in the population, there is a need to adjust the predictions for the criterion shift in each dataset. McManus outlined the approach for applying these corrections within datasets featuring triplets using the following approach.

Matrices P and Q serve as transition matrices, operating on the offspring and the parents, respectively. P is a matrix of the probabilities that offspring measured as right- or left-handed are truly right- or left-handed. Q is a matrix of the probabilities that specific phenotypes observed in mating ( are truly phenotypes (.

Matrix T compiles the predicted probabilities associated with offspring displaying specific phenotypes given the parental phenotypes. This matrix is crafted in accordance with the Mendelian system and the model assumptions regarding the phenotype-genotype system.

Using these three matrices, we can derive the corrected matrix M, with elements representing the predicted probabilities of offspring manifesting as right- or left-handed being born to matings classified as ,

In alignment with McManus' matrix construction methodology for the triplet data, we have extrapolated a parallel correction approach to datasets involving families with multiple children and those involving twins. The comprehensive mathematical derivations underlying the matrix construction are explained in Appendix 1.

## Model fitting

Given the above model and adjustment…… McManus took the following approach to fit the…..:

To estimate the parameters and and evaluate the model's agreement with observed data, a multi-step maximum-likelihood estimation (MLE) method was employed. Initially, a pair of parameter values () was selected. These values were then applied along with the measured rate of left-handedness in each dataset to compute the adjusted predicted distribution for the dataset, M. Subsequently, the log-likelihood was computed using both the original data and the corrected predictions for handedness, M. A detailed explanation of the likelihood function provided in Appendix 2.

The method was executed in two phases. In the first phase, various combinations of within the range of with increments of and within the range with increments of were examined to find the parameter pair with maximum likelihood. This phase provided an approximation to the parameter values that maximize the log-likelihood function.

In the second phase, the parameter pair obtained in the previous phase served as the initial input for a quasi-Newton method, enabling an accurate estimation of the parameter values that maximize the log-likelihood function.

To evaluate how well the MLE parameters and model fitting align with each dataset, as well as with the data as a whole, McManus stated the use of a test statistic that may be treated as statistic. This statistic is calculated by measuring the difference between the level of support indicated by the model and the theoretical level of support expected from a 'perfect fit' of the model to the data, scaled by a factor of 2. This description led us to believe he used the Wilks' likelihood ratio test. In this test, a model is deemed a good fit to the data if the resulting p-value exceeds 0.05.

## Statistical analysis

Upon applying the maximum likelihood estimation across all the datasets McManus found maximum support with a log-likelihood of for the parameter estimates and . (There results were referred to as Model A). For these data he calculated that 'perfect fit' will result in log-likelihood of . The statistical test, indicated a lack of adequate fit to the data. Therefore, McManus concluded that the model was not suitable. He suggested that this outcome might be attributed to the susceptibility of maximum likelihood methods to be influenced by outlier data points.

To identify the outlying datasets, McManus utilized the model parameters with values of and which was referred to by McManus as model C.. He conducted a goodness-of-fit test for each table and dataset individually using these parameters.

In Table 4, presenting the twin data, he calculated an overall , thus he suggested no outlying dataset exists in this table. However, he found that the model failed to explain the DZ twins dataset of Zazzo (1960), and both the MZ twins and DZ twins datasets of Loehlin & Nichols (1976) at a significance level of .

For Table 3, presenting families with multiple offspring, he calculated an overall . Therefore, he suggested that the table doesn't contain outliers. Additionally, he found that the model succeeded in explaining all the datasets within the table.

In Table 2, presenting triplets, he calculated an overall , therefore he suggested that the table contains outliers. Upon examining each dataset individually, he found that the model fails to explain the datasets of Chaurasia & Goswami, Ramaley, Merrell, and McGee & Cozad, each yielding p-values of less than 0.01. Consequently, he labeled them as outliers and tested the goodness of fit of the table after removing these datasets, resulting in .

After identifying the outlying datasets, McManus repeated the MLE process over the data without those datasets and found maximum support of for the estimators , and . He calculated indicating an adequate fit. He referred to this reduced model as **Model B**.

## Results

In our analysis of the complete dataset from tables 2, 3, and 4, which is analogous to **Model A** as proposed by McManus, we obtained a maximum log-likelihood of . This result aligns with parameter values of and . We calculated that for a 'perfect fit', the log-likelihood would be . These outcomes resulted in = , corresponding to a p-value of .

Our results are notably similar to McManus'. Firstly, the parameters we estimated were akin to McManus' up to 2 decimal points, indicating practical similarity. Secondly, McManus' calculated maximum likelihood for Model A is only 3.222 lower than ours, implying a comparable fit between the models. Lastly, as stated by McManus, the results indicated that Model A fails to adequately explain the data.

For the data analogous to the data used in McManus' **Model B**, we obtained maximum log-likelihood of for the estimates of and . Despite the marginal difference in the maximum log-likelihood between our calculation and the log-likelihood reported by McManus, along with the practical similarity of the estimates, we calculated a corresponding to a p-value of . This led us to the conclusion that the model obtained is still not adequate.

We decided to revalidate identifying all outlying datasets, using McManus' method of calculating the goodness of fit for each table and dataset, using the parameters values of and .

Upon examining Table 2, we calculated , Indicating lack of fitness as stated by McManus. When examining each dataset individually, we found that the model fails to explain three datasets out of the four identified by McManus as outliers: Merrell; Ramaley; and McGee & Cozad, each with significance level of less than . However, the dataset of Chaurasia & Goswami seemed to present good fit to the model (. After removing these three datasets, we obtained , correspondung to p-value of , the data fits within the model.

In Table 3, which presents the family data collected by McManus, the overall value received was , implying adequate fit. However, we observed that the model fails to explain the dataset of *ICM2 maternal*, yielding . Explanation to this difference from McManus (1985) presented in the discussion section below.

In Table 4, concerning the twins datasets, the overall value received was indicating a good fit. However, upon closer inspection of the table, we observed that while the monozygotic twins datasets align well within the model ( ), the results for the dizygotic datasets resulted in , corresponding to a p-value of . The lack of fit appears to be influenced by the dataset of Loehlin & Nichols. This dataset exhibited significant values of for both MZ and DZ twins ( respectively) and an overall score of (). After removing this dataset, a score of was obtained for the entire table, and the values obtained for both MZ () and DZ () twins separately indicated adequate fit.

We repeated the MLE process over the data after excluding the 4 outlying datasets. The maximum log-likelihood found was for and which we named Model D. This log-likelihood value is 0.853 higher than the log-likelihood calculated for the parameters derived from the entire dataset (Model A), indicating that the exclusion of the outlier datasets did not significantly shift the estimates. Furthermore, when compared, the parameters of model D and the parameters from McManus' model C showed difference of only 1.222 in log-likelihood. This finding suggests that both models demonstrate comparable fit, implying their near equivalence. Comparison of the models' goodness of fit is presented in Table 5, and Figure 1.

Table 5. Comparison of maximum log-likelihood of the present model and the model of Mcmanus

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Data | | Degrees of freedom | Maximum log-likelihood |  |  |  | Log-likelihood difference from the parameters  from McManus 1985 |
| Fitted total data | Perfect fit (McManus 1985) |  |  | - | - | - | - |
| Model A (McManus 1985) |  |  |  |  |  | - |
| Perfect fit (this study) |  |  | - | - | - | - |
| Model A (this study) |  |  |  |  |  | -3.222 |
| Fitted reduced data (McManus) | Perfect (McManus 1985) |  |  | - | - | - | - |
| Model B (McManus 1985) |  |  |  |  |  | - |
| Perfect fit (this study) |  |  | - | - | - | - |
| Model B (this study) |  |  |  |  |  | -0.765 |
| Fitted reduced data (ours) | Perfect fit (this study) |  |  | - | - | - | - |
| Model D (this study) |  |  |  | 0.067 |  | - |

A close-up of a person's face

Description automatically generated

Figure 1. A contour map of the goodness-of-fit of the models.   
Four points (A, B, C, D) represent the model parameters of each of the four models. Contours represent the difference in support units from the maximum likelihood found.  
 left – likelihood calculated for the complete datasets (i.e., model A).  
 middle - likelihood calculated for the reduced datasets (i.e., model B).  
 right – likelihood calculated for the reduced datasets (i.e., model D)

## Discussion

Up to this point, we've described the steps to reconstruct McManus' 1985 genetic model of handedness. In this section, we will address several questions to measure our success in replicating the model and illuminate new insights regarding the model and its parameters revealed through our work.

Our success in this reproduction attempt was determined by two crucial factors: **the ability to estimate similar model parameters through maximum likelihood estimation and the similarity of our statistical test results with the fit between the model and the observed data.**

Regarding the estimated parameters and maximum likelihood estimation, it's essential to highlight that our results closely aligned with those in McManus (1985). Although, the results for Model A are differ enough to suggest that the discrepancies may be solely attributed to different numeric precision. The reduction in the disparity of maximum likelihoods calculated for Model B compared to Model A between McManus (1985) and our findings suggests that some of the four outlier datasets identified by McManus may contribute to these differences.

Using the 'fitted values' reported by McManus, we reconstructed the model predictions for each of the 4 datasets and evaluated the likelihood McManus should have received for each. While 3 of the 4 datasets showed minor differences between McManus (1985) and our findings, which could be attributed to the improved numerical precision in the present computational methods compared to those available to McManus in 1985, the dataset of Chaurasia and Goswami showed a substantial difference of support units between our results and the expected likelihood basing on McManus report.

Upon inspecting this dataset, the cause for the difference was the disagreement over the measured rate of left- handedness in the parental generation (). McManus measured rate of , while our calculation yielded a rate of . To validate this, we repeated the MLE process using McManus measured rate for this dataset, resulting in maximum likelihood of for the pair of and . These outcomes closely align with those reported in McManus (1985). Therefore, we can confidently declare that we succeeded in estimating the model parameters, and the differences arise from the miscalculation of the rate in McManus (1985) and differences in numeric precision.

In our comparison of the test statistic, we noticed major differences in only 3 out of the 29 datasets when comparing our calculations to those reported by McManus. One of these instances relates to the observations of triplets from Chaurasia and Goswami, which we previously discussed as the cause of the difference.

For the remaining two datasets, we repeated the process using the 'fitted values' reported in McManus (1985) to compute the expected likelihood. Interestingly, for both datasets, the expected likelihood closely resembled our results. Hence, it appears that a miscalculation of the 'perfect fit' occurred for these datasets.

Assuming this to be the cause for the discrepancies, we determined that the test statistics reported in McManus should have been for the Merrell dataset and for ICM2 maternal. These differences from McManus (1985) also suggest that the 'perfect fit' values that should have been reported by McManus were for Model A and for Model B.

After finding those discrepancies of the 3 datasets, we were able to extrapolate the anticipated overall fit of the models that should have been reported by McManus, assuming no significant difference between the goodness of fit for the parameters of Model A, Model B, and Model C, as depicted in figure 1. For model A, which used the entire data, we would have expected the test statistic reported in McManus (1985) to be smaller by 1.781. For model B, utilizing reduced data, we would have anticipated the reported test statistic to be smaller by 2.933 from our calculations. Hence, we can confidently assert that we have successfully replicated the model.

In light of the corrections we made to McManus' statistical analysis results, two significant questions have arisen. Firstly, we need to evaluate whether all the datasets removed in model B should have been excluded. While we acknowledge that removing outlier datasets is a valid method to reassess the model's fitness, we strongly disagree with the decision to remove the dataset of Chaurasia & Goswami. As demonstrated earlier, this dataset was mistakenly classified as an outlier, and our examination of the data revealed that it aligns with the model. Thus, there is no compelling reason to exclude it.

Furthermore, we believe that the datasets chosen for removal in our model D provide a more appropriate choice. Model D not only addresses the issue of outliers and overall fit of the model but also accounts for the problematic fit of the dizygotic twins datasets to the model. This approach ensures more balanced and accurate representation of the data, without excluding datasets that appear to align with the model.

Secondly, we must consider whether the parameters chosen are appropriate to describe the model. Building on the conclusion from the previous question, we can now examine whether the estimators to the parameters selected for the model by McManus (Model C) are appropriate to describe the model.

While both estimations of show a rather similar fit to the datasets when analyzing each dataset alone, and have only a difference of in log-likelihood, we argue that our estimation is more faithful to the results of models A and D. This is because the value of lies between the estimations found in both models. Additionally, while the p-value for the overall fit of the estimates suggested by McManus remains below 0.05 when calculated for the complete dataset and the two reduced datasets, our estimation of yields p-value of for the data used to obtain model D, reinforcing the validity of our estimation.

Finally, we need to address the weaknesses of the model and consider its future implications, continuing the research to enhance its performance and applicability. In addition to the issues raised by McManus in the original paper, we have identified a few more deficiencies. First, as illustrated in figure 1, there is a range of values for the true incidence of left handedness which yields a similar fit between the model and the data, whether we analyze each dataset separately, each table individually, or the fit of the model to the entire dataset. This phenomenon may be attributed to the assumption of uniform distribution of the alleles between all the studies used, because, as presented in equation 1, there is a connection between the allele frequencies, the true incidence of left-handers, and the predicted proportion of left-handers among heterozygotes.

Second, the model relies on data collected almost half a century ago. Since then, social, and demographic changes may have influenced the prevalence of left-handedness. Therefore, it would be worthwhile in the future to revalidate the model using present day data in order to maintain its relevance.

Third, in the time since the publication of the original model in 1985, there have been other models presented that succeeded in fitting more of the datasets than this model. It may be worth comparing McManus' model with these alternative models using additional datasets that represent diverse populations and generations. Such a comparative analysis can provide valuable insights into the processes determining handedness and help us refine our understanding of handedness.  
  
In conclusion, our replication of McManus' 1985 genetic model of handedness has been successful, with parameters closely matching the original model and corrections made to reported statistical values. However, we must acknowledge its weaknesses and the need for future research to enhance its performance and relevance. The model continues to be a valuable tool for understanding handedness, but ongoing updates and refinements are necessary to keep it aligned with modern knowledge and demographic changes.

## APPENDIX 1: CORRECTION MATRICES

In this appendix, we outline the correction matrices used to adjust the predictions of the model to the datasets. Between the datasets involving triplets, families with multiple children, and twins the correction method stays similar, although slight changes are made with the matrices.

### Correcting triplets datasets

For datasets presenting triplets we used the correction matrices as devised by McManus:

P is the transition matrix for the progeny, and it is created using two parameters: t (the true incidence of left-handedness in the population) and p (the observed frequency of left-handedness in the progeny, calculated from the dataset). In creating this matrix, its assumed by McManus that either some left-handers been misclassified as right handers (resulting in ), or that some right-handers been misclassified as left handers (resulting in , but not both. Therefore, the probabilities can be defined as follows:

Q is the transition matrix for the parents, created using two parameters: t (the true incidence of left-handedness in the population) and q (the observed frequency of left-handedness in the parents, calculated from the dataset). Also, the same assumptions about misclassifying handedness as in the creation of P been used. Therefore, Q can be defined as:

Matrix T compiles the anticipated probabilities of offspring displaying specific phenotypes based on the parental phenotypes, following the Mendelian system and the model assumptions regarding the phenotype-genotype system. Detailed calculations are provided in Appendix 2.

Matrix M represents the expected frequencies of children manifesting specific phenotypes, based on the observed parental phenotypes. It is calculated as

Although he didn't state the process of creating the matrices for families with two or more children, McManus did mention that a similar process to the one described above can be used under the assumption that the same transition matrix may be applied independently for each child. Based on this assumption and drawing from the process described by McManus, we have devised methods to calculate corrections for twins and families with multiple offspring.

### Correcting families with multiple children

For datasets featuring families with multiple children, the matrices P and T are adapted to account for scenarios with N children within a family.

For matrix P, originally presented the chances of progeny's handedness being measured as given their true handedness is , a similar approach applied for multiple offspring. In this scenario, P should represent the chances of the handedness of N progenies being measured given their true handedness is . misclassifications occur either as some of the left-handers being misclassified as right-handers or some right-handers being misclassified as left-handers, but not both, the probabilities can be formulated as follows:

The probabilities for each progeny calculated as in P:

In matrix T, the column corresponds to the probabilities of having progenies expressing left-handedness, resulting from the mating of parents with true phenotypes and and producingoffspring. These probabilities are computed using the model parameters, and the comprehensive calculations are provided in Appendix 2.

Similar to the approach used for triplets, we can employ the transition matrices to transform matrix T into matrix M. In matrix M, the column illustrates the anticipated frequencies of having progenies measured as left-handers, based on the observed parental phenotypes.

### Correcting for twins

For datasets involving twins, we use a single transition matrix P, as parental phenotypes are not available.

P is calculated similarly to the case of families with 2 children .

Matrix T is computed differently for monozygotic twins and dizygotic twins, with detailed calculations presented in Appendix 2.

Matrix M, presenting the expected probabilities for encountering twins of phenotype can be received by multiplying T by P.

## APPENDIX 2: PROBABILITIES AND LIKELIHOOD FUNCTIONS

The model proposed by McManus assumes that the genes responsible for determining handedness follow the Mendelian laws of heredity. Under this assumption, along with McManus's suggested probabilities for handedness in homozygote, we were able to create probability functions for the familial data, MZ twins, and DZ twins.

Allele C frequency:

To determine the frequency of allele C in the population, we utilized the model parameters . This determination is based on the expression of the frequency of left-handedness:

Given the genetic system involves only alleles C and D, we deduce , resulting in:

Adding the constraints we receive:

ifand else.

Equation 1. The frequency of allele C in the population

Probabilities of handedness in progeny:

Using Mendel's laws of heredity, we can determine the chances of the resulting offspring from the mating of two genotypes (G1×G2) having each of the genotypes DD, DC, and CC. By using these probabilities, we can find the probabilities of the offspring the present each phenotype:

Subsequently, through the application of the binomial distribution, we can derive the probabilities of progenies within a family of N exhibiting left- handedness:

Thus, given the phenotypes of two parents, the probability of them having n out of N offspring who manifest left-handedness is calculated by:

Equation 2. probabilities of progeny expressing left-handedness given parents handedness

Probabilities of handedness in twins:

In twin data analysis, parental phenotypes are unknown. Dizygotic twins, unlike monozygotic twins, can have different genotypes similar to two non-twin siblings. Given the parents genotypes the probabilities of having twin couple is:

For monozygotic twins, since both twins share the same genotype, Given the parents genotypes the probabilities of having twin couple is:

Therefore, the probability of having twins with phenotypes given the parental phenotypes is:

Thus, the probability of having twin couple with phenotypes is:

Equation 3. probability function of twins expressing phenotypes

### Likelihood functions

To find the best estimates for and , a maximum likelihood estimation approach was used over the datasets in Tables 2,3,4. The probabilities used in the likelihood function obtained from applying correction matrices on the datasets, resulting in different likelihood function for each dataset. Multiplying all the likelihood functions together, the likelihood function of the model was obtained. These functions were also later used for the statistical test.

In Table 2, which presents triplets datasets, using Equation 2 yields the following function for each presented dataset:

Equation 4. likelihood function for triplets

Where represents the observed number of progenies expressing phenotype (right if , left if ) resulting from mating .

In Table 3, which presents families with multiple offspring, using Equation 2 for each dataset yields the following function:

Equation 5. likelihood function for families with multiple offspring

Where represents the observed number of families of size with left-handed progenies resulting from mating .

For table 4, likelihood functions were applied separately for DZ and MZ twins, using the matching probabilities from Equation 3:

Equation 6. likelihood function for twins

Where is the observed number of twins couples where one presents phenotype and the other presents phenotype .























## Bibliography

Annett, M. (1972). The distribution of manual asymmetry. *British Journal of Psychology*, *63*(3), 343–358. https://doi.org/10.1111/j.2044-8295.1972.tb01282.x

[Conard, N. J. (2011). *Neanderthal Lifeways, Subsistence and Technology: One Hundred Fifty Years of Neanderthal Study (Vertebrate Paleobiology and Paleoanthropology)* (2011th ed., p. 670). Springer.](https://sciwheel.com/work/bibliography/15365201)

[Coren, S., & Porac, C. (1977). Fifty centuries of right-handedness: the historical record. *Science*, *198*(4317), 631–632. https://doi.org/10.1126/science.335510](https://sciwheel.com/work/bibliography/10047819)

[McManus, C. (2019). Half a century of handedness research: Myths, truths; fictions, facts; backwards, but mostly forwards. *Brain and Neuroscience Advances*, *3*, 2398212818820513.https://doi.org/10.1177/2398212818820513](https://sciwheel.com/work/bibliography/14962387)

[McManus, I. C. (1985). Handedness, language dominance and aphasia: a genetic model. *Psychological Medicine. Monograph Supplement*, *8*, 1–40. https://doi.org/10.1017/S0264180100001879](https://sciwheel.com/work/bibliography/15020506)

[Porac, C. (2016). *Laterality: Exploring the Enigma of Left-Handedness* (1st ed., p. 232). Academic Press.](https://sciwheel.com/work/bibliography/15365161)