

ABSTRACT

Title of Dissertation: LIFE'S RICH PATTERN: THE ROLE OF STATISTICS
AND PROBABILITY IN NINETEENTH CENTURY
ARGUMENTATION FOR THEORIES OF
EVOLUTION, VARIATION, AND HEREDITY

James Ian Wynn, Doctor of Philosophy, 2006

Dissertation Directed By: Jeanne Fahnestock

Department of English

Though modern philosophers of science recognize the inappropriateness of the reduction of all scientific investigations to mathematics, mathematics and science share a long history with one another during which mathematics has been employed as a major component of scientific argumentation. Over the last twenty years, rhetoricians have done substantial work studying the role of argumentation in science (Bazerman 1988; Gross 1990, 2002; Myers 1990; Fahnestock 1999); however, despite the importance of mathematics in making scientific arguments, little effort has been made to understand the role mathematics has played in making these arguments. This dissertation represents a move to resolve this shortcoming by investigating the role of mathematics in arguments in evolutionary biology from the middle of the nineteenth to the beginning of the twentieth century.

In the first part of the nineteenth century, the mass collection and mathematical assessment of data for scientific purposes provides the context for understanding some of the rhetorical choices of an important group of natural philosophers and biologists who developed arguments in the second half of the century about the nature of variation, evolution, and heredity. In the works of Charles Darwin, Gregor Mendel, Francis Galton, and Karl Pearson, arguments from probability and statistics play important roles as support for their arguments and as a source of invention for their claims.

This investigation of the rhetorical situations of these four biologists, their arguments, and the role of the principles, operations, and formulae of probability and statistics supports the position that mathematization had a major impact on the nature of scientific evidence in the nineteenth century. What it also suggests is that, though mathematized arguments may have had a great deal of credibility within the scientific community in general, factors such as the stature of the rhetor and of their biological theory within their specific discourse communities played an equally important role in the persuasiveness of their arguments.

LIFE'S RICH PATTERN: THE ROLE OF STATISTICS AND PROBABILITY IN
NINETEENTH CENTURY ARGUMENTATION FOR THEORIES OF EVOLUTION,
VARIATION, AND HEREDITY

by

James Wynn

Dissertation Submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements of the degree of
Doctor of Philosophy

2006

Advisory Committee:

Professor Jeanne Fahnestock, Chair
Professor Gerald Borgia
Professor Jane Donawerth
Professor Michael Israel
Professor Shirley Logan

©Copyright by James Ian Wynn

2006

DEDICATION

To my wife Gina, daughter Sofia, and son Alexander

TABLE OF CONTENTS

LIST OF FIGURES.....	v
Introduction	1
Chapter 1: Arithmetic of the Species: Darwin and the Use of Mathematics in	
Argumentation	6
Mathematics and Argument	7
Mathematical Darwin?	8
Mathematics in <i>The Origin of Species</i>	9
Variation under Domestication	14
Variation under Nature	18
The Struggle for Existence	25
Natural selection: calculating diversity	37
Conclusion	44
Chapter 2: Drawing Back the Veil: Probability and Statistics from 1600 to 1900 ...	46
The Seventeenth Century: Pascal, Fermat, and Huygens	47
Jakob Bernoulli and Abraham de Moivre	50
Pierre-Simone Laplace	55
Probability and Statistics in the Nineteenth Century	60
Conclusion	69
Chapter 3: Hidden Value: Mendel, Mathematics, and the Case for Uniform	
Particulate Inheritance	71
The Hybridist Tradition	74
Joseph Gottlieb Kölreuter	75
Carl Friedrich von Gärtner	81
The Origin of Mendel’s Experiments	85
Mendel the Mathematician and Physical Scientist	93
The Role of Quantification, Combinatorics, and the Principles of Probability in	
Making Mendel’s Case	95
From Word into Law	106
Critical Responses to Mendel’s “Experiments in Plant Hybridization”	124
Conclusion	132
Chapter 4: Contours of Heredity: The Law of Error and Francis Galton’s	
Biometrical Approach to Heredity	135
Francis Galton: a Gentleman of Science	136
The Development of a Mathematical Theory of Inheritance	142
<i>Natural Inheritance</i>	149
Chapter One: The Author, the Audience, and the Journey	150
Chapters Two and Three: Establishing Jurisdiction	153
Chapters Four and Five: Demonstrating Compatibility with the Law of Error . 164	
Making the Case for Regression and the Geometric Diminution of Variation:	
Blended Inheritance.	172
Predicting Regression: Exclusive Inheritance	182
The Reception of <i>Natural Inheritance</i>	188
Conclusion	192

Chapter 5: Behind the Curve: The Place of Mathematics in Making Claims about Heredity at the Opening of the Twentieth Century	195
Karl Pearson (1875-92).....	196
On the Road to Biometry (1893-96)	204
Weldon and Pearson.....	205
Oil and Water?	209
The Biometrik Response	215
Homotyposis and Heredity.....	220
Bateson's Critique	233
Conclusion	247
Epilogue	249
Appendix A	255
Appendix B	257
Appendix C.....	258
Appendix D.....	260
Appendix E	261
Appendix F	263
Appendix G.....	264
Appendix H.....	265
Appendix I	266
Appendix J.....	267
Appendix K.....	268
Appendix L	270
Appendix M.....	272
Appendix N.....	275
References.....	277
Works Consulted.....	283

LIST OF FIGURES

<i>Fig. 1. Regression from Mid-Parent to Son.....</i>	<i>174</i>
<i>Fig. 2. Geometrical Proof of the Filial Regression Value from Natural Inheritance.....</i>	<i>176</i>
<i>Fig. 3. Geometrical Proof of the Filial Regression Value with the Actual Data.....</i>	<i>177</i>

Introduction

*I assert... that in any special doctrine of nature there can be only as much proper science as there is mathematics therein. For... proper science, and above all proper natural science, requires a pure part lying at the basis of the empirical part.—Kant in *Metaphysical Foundations of Science*¹*

Though modern philosophers of science recognize the inappropriateness of the reduction of all scientific investigations to mathematics, mathematics and science share a long history with one another during which mathematics has been employed as a major component of scientific argumentation. Over the last twenty years, rhetoricians have done substantial work studying the role of argumentation in science (Bazerman 1988; Gross 1990, 2002; Myers 1990; Fahnestock 1999); however, despite the importance of mathematics in making scientific arguments, little effort has been made to understand the role mathematics has played in making these arguments. This work represents a move to resolve this shortcoming by investigating the role of mathematics in arguments in evolutionary biology from the middle of the nineteenth to the beginning of the twentieth century.

In *Science in Culture: The Early Victorian Period*, Susan Cannon argues that in the opening half of the nineteenth century the sciences were transformed by two fundamental changes in the way that natural investigations were conducted. First, unlike the previous century, there was a greater drive to collect data of all kinds. This urge resulted in what historian Ian Hacking calls an “avalanche of numbers” (Hacking 5). Data on medical conditions, suicides, crimes, marriages, weather, geology, electricity,

¹ Immanuel Kant, *Metaphysical Foundations of Natural Science*, ed. Michael Friedman (Cambridge: Cambridge UP, 2004) 6.

magnetism, etc. were called for by scientists, politicians, and tradesmen and collected as interest in these phenomena developed (Cannon 225-234).

Second, in addition to an increased drive to collect data, there was also a change in the manner of treating evidence. The first half of the nineteenth century witnessed the rise of the mathematically equipped scientist. Whereas the mathematical sciences had been confined to astronomy and geodesy in the eighteenth century, in the beginning of the nineteenth century mathematics began to expand to the study of magnetism, electricity, fluids, and even to sociology. This expansion, along with the reverence accorded to the already established mathematical sciences, suggests a growing belief among scientists and natural philosophers that mathematical arguments represented the gold standard for describing and reasoning about natural phenomena (Cannon 234).

The mass collection and mathematical assessment of data for scientific purposes which began in the first part of the nineteenth century provides the context for understanding some of the rhetorical choices of an important group of natural philosophers and biologists in the second half of the century who developed arguments about the nature of variation, evolution, and heredity. In the works of Charles Darwin, Gregor Mendel, Francis Galton, and Karl Pearson, arguments from probability and statistics play important roles, from ethical support for their arguments to sources of invention for claims.

In addition to showing the importance of these mathematical fields as resources for developing ethos and argument, the use of mathematics in these biological arguments also reveals the limitations which these types of arguments face. Despite the generally accepted belief among scientists that mathematics was a powerful ally, there were also

disagreements over what could be treated mathematically and who was qualified to speak on these issues. In certain cases, challenges trumped arguments invented from or supported by mathematical principles, operations, or formulae.

The goal of this work is to investigate the roles of statistics and probability in making arguments about variation, evolution, and heredity in the late nineteenth century. In chapter one, “Arithmetic of the Species,” the role of quantification and arithmetical operations in Darwin’s arguments in the *Origin of the Species* will be examined. The case will be made that, though the *Origin* is not overtly mathematical, closer inspection of the arguments in the text and other writings of Darwin will reveal the degree to which he employs quantification to add an ethos of precision to his arguments. In addition, it will also discuss the importance of arithmetical operations in helping him discover the principle of divergence of character.

The second chapter, “Drawing Back the Veil,” presents a brief history of probability and statistics for the purpose of explaining some of the mathematical concepts that come to play a prominent role in the development of biological arguments in the latter half of the nineteenth century. In addition, it describes in greater detail the socio-historical context in which these ideas and practices evolved to become a central part of the debate surrounding these biological phenomena.

Chapter three, “Hidden Value,” examines the arguments of Gregor Mendel in his groundbreaking work in genetics, “Experiments in Plant Hybridization.” It investigates the role of the principles, operations, and formulae of combinatorics, probability, and statistical thinking in making his case for the regularity of the distribution of characters over generations, the independent segregation and integration of character pairs, and the

independent and equal contributions of parental characters in peas. Additionally, this chapter explores the reasons for the failure of these mathematically informed arguments to find an audience despite their mathematical rigor.

Francis Galton and his arguments for a theory of variation and heredity based on the law of errors in his seminal work *Natural Inheritance* is the subject of the fourth chapter, “Contours of Heredity.” This chapter examines the rhetorical factors that contribute to Galton’s success in approaching variation and heredity with the principles, operations, and formulae of the law of errors and statistics. It also examines his arguments and explains how the law of error serves as an organizing analogy for his description of the hereditary process and reproductive outcomes. The case is made here that the success of Galton’s persuasive campaign can be attributed to multiple factors including his ethos, his relatively orthodox position on heredity, and his use of an overt campaign to persuade his readers to accept the application of the law of error to biological phenomena. It will also be argued that Galton’s *Natural Inheritance* laid the groundwork for the resurgence of Mendel at the turn of the century.

The final chapter, “Behind the Curve,” explores the arguments of the mathematician Karl Pearson in his paper “On the Principle of Homotyposis and its relation to Heredity.” The central focus of this chapter is to understand to what degree mathematics was accepted by biologists in the last decade of the nineteenth century and to what degree the acceptability of a mathematical line of argument is contingent on the acceptability of a biological line of argumentation. In order to answer these questions, I explore the rhetorical context in which the mathematical arguments in Pearson’s

Homotyposis paper are made and his theoretical assumptions about variation and heredity.

This investigation of the rhetorical situations of these four biologists, their arguments, and the role of the principles, operations, and formulae of probability and statistics supports the notion suggested by Cannon that mathematization had a major impact on the nature of scientific evidence in the nineteenth century. What it also suggests, however, is that, though mathematized arguments may have had a great deal of credibility within the scientific community in general, factors such as the stature of the rhetor and of their biological theory within their specific discourse communities played an equally important role in the persuasiveness of their arguments.

Chapter 1: Arithmetic of the Species: Darwin and the Use of Mathematics in Argumentation

During the three years which I spent at Cambridge my time was wasted, as far as academic studies were concerned.... I attempted mathematics, and even went during the summer of 1828 with a private tutor... I have deeply regretted that I did not proceed far enough at least to understand something of the great and leading principles of mathematics.... But I do not believe I would have succeeded beyond a very low grade.-
Charles Darwin in his *Autobiography*²

*The characteristic bent of Charles Darwin's mind led him to establish the theory of descent without mathematical conceptions; even so Faraday's mind worked in the case of electromagnetism. But as every idea of Faraday allows of mathematical definition, and demands mathematical analysis in its modern statement, so every idea of Darwin—variation, natural selection, sexual selection, inheritance, prepotency, reversion—seems at once to fit itself to mathematical definition and to demand statistical analysis. –Karl Pearson in *Biometrika*³*

In order to follow the development of the role of statistics and probability in making arguments about evolution, variation, and heredity, it is important to start by looking at the man whose arguments provided not only theoretical foundations for the modern study of evolution and variation, but also inspiration for modern mathematical approaches to these subjects. That man, of course, is Charles Darwin whose seminal work *The Origin of the Species*, although not the first text to suggest evolution, was the first to present in a systematic and well-supported fashion the basic tenets of evolution including the all-important mechanism of natural selection.

The purpose of this chapter is to explore the role that precise quantification and mathematical operations and principles play in the development and support of Darwin's position in *The Origin of Species* as well as to establish a starting point from which to

² Charles Darwin, *Autobiographies*, eds. Michael Neve and Sharon Messenger (London: Penguin Books. 1986) 30.

³ Karl Pearson, "The Spirit of Biometrika," *Biometrika* 1.1 (1901) 3-4.

understand the larger developments in the mathematization of arguments in evolution, variation, and heredity from the middle of the nineteenth to the beginning of the twentieth century. To accomplish these goals, I will argue, contrary to most current scholarship, that Darwin employs mathematically informed arguments to support his conclusions on variation in domestic and natural stocks of organisms, the existence of competition between organisms for resources, and the importance of diversity in succeeding in the struggle for existence. In these arguments, Darwin relies on mathematical symbols, operations, and principles to supply the evidence and, in some cases, the understanding of nature needed to build his case for evolution and natural selection.

Mathematics and Argument

Before discussing the specific mathematical features and practices that Darwin employs in supporting and inventing his biological arguments, it is first important to understand clearly what is intended here by the term “mathematics” as well as what roles mathematics can play in making arguments. Admittedly, “mathematics” is broad and fairly ambiguous term referring to a whole host of symbols, operations, formulae, and principles designed to represent quantities or magnitudes and their operations, interrelations, combinations, generalizations and abstractions.⁴ For the purposes of this investigation, however, the concept of mathematics will be limited to the particular symbols, operations, formulae, and principles from fields of arithmetic, statistics, probability, and combinatorics that are used to make arguments about variation, evolution, and heredity in the latter half of the nineteenth century.

⁴ “Mathematics,” *Merriam-Webster’s Collegiate Dictionary*, 10th ed. 1998.

When making an argument, mathematical symbols, operations, formulae, and principles can play two distinct roles. On the one hand, they can be used as *support* for non-mathematical conclusions or traditionally qualitative lines of argument either by adding precision or by adding an air of rigor or deductive certitude to arguments. On the other hand, mathematical operations and principles can be used to *discover* or *invent* new lines of argument. New arguments can be discovered when mathematical principles and operations guide the rhetor to adopt a particular qualitative argument. They can be invented when mathematical principles themselves serve as the source for a new line of argument which does not exist in more traditional lines of non-mathematical argument.

Mathematical Darwin?

Although Pearson, in the quotation above, recognizes a fundamental mathematical inspiration in Darwin's theories, Darwin scholars who have investigated *The Origin of Species* have either ignored or mostly dismissed the idea that mathematics plays any significant role in his efforts to make the case for evolution and natural selection. A survey of 17 books and articles on this subject, most of which were published in the last twenty five years,⁵ reveals the uniformity of this scholarly dismissal. Among the book-length works, few texts make any association between mathematics and Darwin's arguments and those that do (Gale 1982, Ghislen 1969, and Hull 1973) predominantly comment either on the lack of mathematical reasoning in the text or on Darwin's inability

⁵ See Appendix A for a full listing of the books used in the survey and the results of the survey for each text.

to use mathematics even if he had wanted to. A review of six journal articles⁶ on the development of Darwin's arguments also reveals a lack of discussion of the role of mathematics in *The Origin of Species*. Of the six journal articles surveyed, only two mention mathematical reasoning at all. Of those two, only one (Browne 1980) clearly asserts the value of mathematical reasoning to Darwin's arguments.⁷

Mathematics in *The Origin of Species*

An examination of selected books and articles which discuss Darwin's arguments provides statistical evidence that the majority of modern Darwin scholars do not believe that mathematics plays an important role in supporting or inventing arguments in *The Origin of Species*. These statistical results raise the question: "If mathematics plays such an important role in *The Origin* why is it that so few scholars have bothered to write about it?" There are, I believe, several reasons for this. First, a cursory review of the text itself reveals that there are very few places where mathematical symbols, numbers, tables, equations, etc. are used. This scarcity of mathematical notation is puzzling even to those who argue in favor of the importance of mathematics in *The Origin* like Janet Browne,⁸ who comments on the scarcity of mathematics in the text.

⁶ See Appendix B for a full listing of the journals used in the survey and the results of the survey for each text.

⁷ All told, out of 17 texts specifically devoted to discussing the logical contents and/or development of Darwin's argument, only three clearly suggest that mathematics has any significant role to play (Browne 1980, Ghislen 1969, and Bowler 1990). That works out to about roughly 18% of texts recognizing the importance of the role of mathematics to Darwin's argument, 11% dismissing it, and 71% ignoring it altogether.

⁸ Janet Browne is a leading Darwin scholar and biographer who has written extensively on Darwin and his work.

That Darwin's botanical arithmetic has been neglected by historians is partly his own fault. In *On the Origin of Species* he barely referred to his botanical statistics or the long sequence of calculations which he had undertaken from 1854 to 1858. He compressed and simplified these into a few meager paragraphs, giving his reader only six pages of statistical data to fill out the discussion of "variation of nature" in Chapter II. (53).

Despite its absence in the actual text, a brief review of selected materials including Darwin's notebooks, the published manuscript of his "big species book" and the first two chapters of that manuscript, which were published separately in 1868 under the title *The Variation of Plants and Animals Under Domestication*, reveal the extent to which mathematics influenced the development of his theories. In these publications, Darwin supplies his readers not only with lists of quantitative evidence and calculations, but also with occasional glimpses of the degree to which these data and calculations helped him formulate his conclusions about natural selection and evolution.⁹

The existence of mathematical symbols and operations in these extrinsic sources, however, still does not explain why, if they were so important to Darwin's argument, the majority of them were left out of his text. The answer to this query is provided by Darwin himself in the introduction of *The Origin*.

I can here give only the general conclusions at which I have arrived, with a few facts in illustration, but which, I hope, in most cases will suffice. No one can feel

⁹ See Appendices C, D, and E. These examples are explained in detail later in the chapter.

more sensible than I do of the necessity of hereafter publishing in detail all the facts, with references, on which my conclusions have been grounded; and I hope in a future work to do this. (4)

Here, Darwin explains that he is able in the text to give only a general outline of his theory and as a result has to forgo presenting all of the data and discussion that he might have otherwise provided. The reason for this brevity is that he has been rushed into publication as the result of the emergence of Alfred Russell Wallace's theory of evolution which, for all intents and purposes, offered the same conclusions as his own. In addition, Darwin's lack of specific, quantitative detail may have been a strategy to make his work accessible to a wider readership for whom a dense data-rich text would have seemed too formidable.

Besides the infrequency of mathematical symbols and operations in *The Origin*, critics such as David Hull argue that Darwin could never have employed mathematics to model or predict the actions of evolution and natural selection even if he had wanted to. In his book *Darwin and His Critics*, Hull takes the position that because of the subject of his research – unpredictable forces that varied over time – there was no available mathematics to aid him in prosecuting his argument:

Darwin could not help but know the crucial role which mathematics had played in physics, since Herschel had repeatedly emphasized it in his *Discourse*, but it did not seem to be in the least useful in his own work in biology.... For Darwin, mathematics consisted of deductive reasoning, and he distrusted greatly

“deductive reasoning in the mixed sciences.” In his own work, he seldom was presented with a situation in which he could use such deductive reasoning. He was constantly forced to deal in probabilities, and no one could tell him how to compute and combine such probabilities. (12-13)

Here Hull reasons that because the forces with which Darwin was dealing were so complex, they could not possibly have been handled by the mathematics which existed at the time. Although basic mathematic principles of probability and statistics did exist at the time Darwin was working, Hull is reasonably justified in assuming that the mathematized statistics needed for modeling the spread of variation through populations of organisms, which would become the hallmark of population genetics in the 1930s, had not yet been developed. The fundamental problem with this position, however, is that Darwin’s goal was not to predict the probability or model the detailed spread of variations through populations of organisms. Instead, the aim of his argument was to prove that in the past, organisms had varied over time and that the force guiding the spread of variation through populations was natural selection. If we understand Darwin’s work with regard to these goals, the sophistication of the mathematics necessary to support and discover arguments is drastically lowered.

In addition to the fact that Darwin had the ability to use mathematical resources in his argument, he would want to present as rigorous an argument as possible for what he realizes is a controversial position. Hull recognizes Darwin’s interest in increasing the rigor of his arguments.

Darwin criticized Lamarck and Chambers, not for suggesting mechanisms of evolution which he thought were mistaken, but for foisting their views on the scientific community without sufficient effort at careful formulation and verification. Lamarck and Chambers looked upon the process of scientific verification as a very casual affair. Darwin looked upon these matters as of utmost gravity.... Darwin emphasized the importance of providing empirical evidence for scientific hypotheses (10-11).

Given that mathematics was becoming an increasingly important feature of scientific argumentation in the physical sciences, if Darwin really were trying to be more rigorous in his arguments, then using precisely quantified evidence and employing mathematical principles and operations to discover arguments would be obvious means of making his arguments more disciplined in the mind of “scientific” audiences. This I contend is exactly what Darwin is attempting to do in developing his arguments for *The Origin*.

A brief examination of the context, arguments, and motives of *The Origin of Species* suggests that Darwin could make his arguments using mathematical principles, symbols, and operations, despite his own admitted lack of mathematical genius. Though these types of arguments are admittedly not prominent in the text, their presence behind the scenes as a driving force for Darwin’s arguments is palpable and will be made clear in the following sections through a detailed analysis of Darwin’s arguments in the first four chapters of *The Origin*.

Variation under Domestication

In the first chapter of *The Origin*, “Variation Under Domestication,” Darwin begins to carefully construct his arguments in support of evolution and natural selection by providing evidence which suggests 1) that organisms are highly plastic and can be made to vary to a great degree and 2) that this variation accumulates over time resulting in populations of organisms which were once related becoming physically distinct. In order to ease his readers into the critical evidence and arguments he will make later in the text, Darwin begins by approaching these issues in terms of a familiar process, domestic breeding. Though some of the persuasive strength of the argument in the first chapter comes from the breeding analogy, mathematical calculation from precisely quantified empirical evidence also supplies necessary support for the case that not only does variation exist in domestic breeds, but also that variations can be accumulated over time through breeding.

One of the major hurdles that Darwin must overcome in order to convince his readers that species evolve over time is that he has to explain how organisms which were once similar to one another could become physically distinct. His strategy in the first chapter to overcome this obstacle is to present his reader with a case of domesticated organisms which reveals how much a species can be varied through human intervention. Darwin chooses a population of organisms which was popular for breeding and which boasted many distinct varieties: pigeons.

After a general introduction laying out the various positions on the degree and time of occurrence of variation in domestic breeds of animals and plants, Darwin

discusses the spectacular physical differences that exist between certain breeds of pigeons.

In the skeletons of several breeds, the development of the bones in the face in length and breadth and curvature differs enormously. The shape, as well as the breadth and length of the ramus of the lower jaw, varies in a highly remarkable manner. The number of the caudal and sacral vertebrae vary as does the number of ribs. (20)

Although no specific quantities are identified in Darwin's assessment of the differences between varieties of pigeons in this passage, claims like, "The number of the caudal and sacral vertebrae vary," suggest that the evidence, if it is to be at all reliable, had to have been derived from the actual counting and comparison of vertebrae in skeletal samples from the different breeds of pigeons. In fact, Darwin had indeed made these measurements but published them later in *The Variation of Plants and Animals Under Domestication*.¹⁰ A series of tables in this text provide the reader with measurements not only of the number of caudal and sacral vertebrae, but also of the average difference between the length of the feet of different species of pigeons and the proportional difference between the length of the feet and the size of their bodies.

Darwin's quantitatively supported argument of the degree to which pigeons vary is an important first step in making his case for variation because it reveals the magnitude of change that is possible in organisms. The next step is to show that existing pigeon varieties are related to a common progenitor and have developed differently over time as

¹⁰ See appendix C for the measurement charts.

a result of active human selection. In making this case, Darwin challenges the primary theory of natural diversity at the time: special creation. Adherents to the theory of special creation held that variation in organisms could be explained not in terms of the transformation of a single distinct population of organisms into a multiplicity of varieties, but by the existence of wild stocks of organisms which accounted for each of the known varieties. In order to cast doubt on this position, Darwin relies on a comparison between the number of wild stocks and the number of domestic varieties of pigeons:

If several breeds are not varieties, and have not descended from the rock-pigeon, they must have descended from at least seven or eight aboriginal stocks.... But besides *C. livia*, with its geographical sub-species, only two or three other species of rock pigeon are known; and these have not any of the characteristics of the domestic breeds. (21)

A quantitative comparison of the number of wild varieties and the number of domestic breeds reveals that there is no one-to-one correlation between the number of domestic and wild varieties. This discrepancy casts doubt on the special creationist position of the fixity of organic forms. Darwin's theory of evolution, however, is able to account for the difference. It suggests that if we allow that populations of organisms vary over time, there is no need for the existence of several feral progenitors, only the existence of a series¹¹ of related forms that reveals the incremental changes in pigeons.

¹¹ In arguing for a series, i.e. an incrementum argument, Darwin is presenting his reader with what Jeanne Fahnestock argues in *Rhetorical Figures in Science* is quintessentially a quantitative scalar argument of the more and the less: "As its first essential feature, an incrementum has to be formed according to some

In *The Origin* Darwin argues that such series of forms do exist among the domestic breeds and that this series can be visually evidenced by juxtaposing closely related pigeons at different levels of the taxonomy with one another:

Although an English carrier and a short-faced tumbler differ immensely in certain characters from the rock-pigeon, yet by comparing the several sub-breeds of these varieties, more especially those brought from distant countries, we can make an almost perfect series between the extremes of structure. (24)

In this opening salvo of his argument, Darwin relies on quantitative comparisons in order to make the tentative case that selection provides a better accounting of the empirical evidence presented by an investigation of breeding practices than special creation. First, he relies on descriptive arguments supported by quantitative comparisons to make the case that selective breeding has been able to create significant variations within and among domesticated pigeon breeds. Next, he quantitatively compares the number of domestic and wild stocks of pigeons to support his contention that variation from a single wild progenitor has led to the current degree of diversity among domesticated pigeons. Finally, he argues for the existence of a series of forms, an argument which is based fundamentally on a quantifiable increase or decrease of some character. Although these traditional lines of comparative reasoning are often posed in natural language in *The Origin* or in terms of qualitative features, they are fundamentally quantitative comparisons which are supported in some cases by quantitative data.

principle of ordering, and by far the most common principle of ordering is by increase or decrease in some quantifier or attribute” (Fahnestock 95).

Variation under Nature

In the second chapter of *The Origin*, “Variation under Nature,” Darwin turns his discussion from the effects of selection on domestic species to the possibility of selection without human intervention. His goal in this chapter is to provide arguments and evidence that variation is a natural process that can spread without human intervention from isolated individual organisms to large populations through the process of natural selection. He achieves this goal with the help of quantitative comparisons and mathematical calculations, which prove that not only are the taxonomic categories found in catalogues of species fuzzy, but also that the source of this fuzziness can be accounted for by conceiving of diversity as the result of the dynamic process of continual variation.

In the first portion of chapter two, Darwin sets up his argument by refuting the position of special creationists who believed that each species identified in a taxonomic hierarchy marked a unique creation which was readily identifiable by the existence of an indelible set of features. He argues that if this position were true, then there should be a definite consensus among naturalists about which organisms belonged in which particular category. In order to test the veracity of this assumption, Darwin undertakes a quantitative comparison of the categorizations made by several experts in the field:

Compare the several floras of Great Britain, of France, or of the United States, drawn up by different botanists, and see what a surprising number of forms have been ranked by one botanist as a good species, and by another as mere varieties.

Mr. H.C. Watson, to whom I lie under deep obligation for assistance of all kinds, has marked for me 182 British plants, which are generally considered as varieties, but which have all been ranked by botanists as species.... Under genera, including the most polymorphic forms, Mr. Babington gives 251 species, whereas Mr. Bentham gives only 112, - a difference of 139 doubtful forms! (41)

Through this comparison, Darwin reveals that even those researchers who claim a high level of expertise in plant identification and categorization come to astonishingly little agreement concerning which organisms should be ranked as varieties within a species and which as separate species. By doing a quantitative analysis here, Darwin provides evidence, using the naturalists' own numbers, that the boundaries between species, varieties, and genera are not as well defined as they believe. By casting doubt on the ability of the special creationist theory of clearly defined taxonomic categories to account for the empirical realities of nature, he creates the opportunity to present his own theories of evolution and natural selection, which he believes more adequately account for the data.

He opens the second portion of the chapter by clearly laying out for his readers his position:

Hence I look at individual differences, though of small interest to the systematist, as of high importance for us, as being the first step towards such slight varieties.... And I look at varieties which are in any degree more distinct and permanent, as steps leading to more strongly marked and more permanent

varieties, and at these latter, as leading to sub-species and to species.... I attribute the passage of a variety, from a state in which it differs very slightly from its parent to one in which it differs more, to the action of natural selection in accumulating... differences of structure in certain definite directions. (44).

In these lines, Darwin presents his reader with a vision of diversity in nature as a dynamic process rather than a static condition. He argues that the small differences we observe in individual organisms can spread throughout successive generations making the offspring of those individuals slightly different from the general population from which they originated. These differences can widen through the continued accumulation of variation and eventually transform distinct varieties into distinct species. This whole dynamic process, Darwin argues, can be attributed to natural selection which he defines as, “the preservation of favorable variations and the rejection of injurious variations” (68).

Once he has established his position on the source and character of diversity, Darwin presents his readers with quantitative evidence to support it. Using available botanical compendiums, he calculates the number of botanical varieties belonging to species with the greatest estimated populations of organisms in hopes of discovering some general pattern in or connection between these species: “I thought that some interesting results might be obtained in regard to the nature and relations of the species which vary most, by tabulating all the varieties in several well-worked floras” (45). The

results of his tabulation reveal that there is a correlation between the size and range of a species' population and the number of varieties recorded for that species.¹²

In any limited country, the species which are most common, that is abound most in individuals, and the species which are most widely diffused within their own country... often give rise to varieties sufficiently well-marked to have been recorded in botanical works. Hence it is the most flourishing, or, as they may be called, the dominant species... which oftenest produce well-marked varieties. (45-46)

What Darwin discovers, or confirms, as the result of his calculations, is that the more populous species tend to have a greater number of identified varieties associated with them. This correlation is accounted for by his dynamic theory of natural diversity because statistically such a correlation between the size of a population and the development of sub-populations would be expected as larger populations would have a greater number of variations which might be spread by natural selection. Darwin in fact suggests this correlation in the first chapter of the text when he writes, "I must now say a few words on the circumstances favorable or reverse, to man's power of selection. A high degree of variability is obviously favorable, as freely giving the materials for selection to work on" (35).

Although the correlation between the size of a species and the number of varieties connected with it seems to support Darwin's argument that diversity in nature is the result of the spread of variations, he recognizes that more data must be presented if he is to

¹² See Appendix D for Darwin's mathematical tables

prove that this relationship is a general fact of nature. Following his examination of species and varieties, he conducts further calculations to assess whether or not the principle holds true for the relationship between genera and species.

If the plants inhabiting a country and described in any Flora be divided into two equal masses, all those in the larger genera being placed on one side, and all those in the smaller genera on the other side, a somewhat larger number of the very common and much diffused or dominant species will be found on the side of the larger genera. (46)

Here, Darwin reaffirms that the same quantitative correlation which exists between species and varieties also exists between genera and species. His calculations reveal that the larger genera, those with five or more species, had species with a greater number of varieties.

Strengthened by the accumulating evidence, Darwin makes a point to emphasize the success of his theory in accounting for the patterns revealed by his calculations:

From looking at species as only strongly-marked and well-defined varieties, I was led to anticipate that the species of the larger genera in each country would oftener present varieties, than the species of the smaller genera; for wherever many closely related species... have been formed, many varieties or insipient species ought, as a general rule, to be now forming.... On the other hand, if we look at each species as a special act of creation, there is no apparent reason why

more varieties should occur in a group having many species, than in one having few. (47)

In addition to playing up the success of his prediction, Darwin also challenges the special creationist position to account for the same results. If the different taxonomic categories did in fact represent unique populations of organisms that shared no relationship with other populations, then what would account for the correlations which his calculations reveal? Though opponents of his theory might argue that these correlations are random coincidence, Darwin suggests here that the fit between the patterns he describes in the quantitative data and the process of variation that he proposes in his theory is too good to be coincidental.

Not sufficiently satisfied with the scope of the evidence he has presented and analyzed thus far, Darwin ends this portion of his argument with the discussion of an even larger data set to ensure that his claims concerning the flora of England are not simply a unique case. In the final data set, he expands his evidentiary support from calculations of ratios of genera, species, and varieties in one country to an analysis of ratios between these different taxonomic levels in 12 different countries:¹³

To test the truth of this anticipation I have arranged the plants of twelve countries, and the coleopterous insects of two districts, into nearly two equal masses, the species of the larger genera on one side, and those of the smaller genera on the other side, and it has invariably proved to be the case that a larger proportion of the species on the side of the larger genera present varieties than on the side of the

¹³ See Appendix E for Darwin's mathematical tables

smaller genera. Moreover, the species of the larger genera which present any varieties, invariably present a larger average number of varieties than do the species of the small genera. (47)

As with the plants of England, Darwin finds in the plants and insects in 12 different countries a correlation between the ratios of population sizes and the amount of variation in sub-populations at different levels of the taxonomic hierarchy. This evidence supports his hypothesis that diversity in nature is a reflection of a dynamic process of variation driven by natural selection and challenges the existing paradigm of special creation because that theory cannot account for this recurrent pattern of correlation.

An analysis of Darwin's argument in the second chapter thus shows the importance of quantitative comparison and other mathematical operations in supporting his theories of evolution and natural selection. In the opening portion of the chapter, he uses quantitative comparison with precise numerical values to challenge the veracity of the existing paradigm of special creation by revealing that among experts there is no clear consensus on the categorization of organisms in nature. He uses this discrepancy as an opportunity to present his own theory, which, he argues, also accounts for the difficulties in developing taxonomies.

Once he has cast doubt on the theory of his opponents and offered his own, he then accumulates evidence to support it. With the help of a previously calculated ratio, he concludes that patterns of correlation exist between the size of populations and the development of variation within related sub-populations. These connections between the sizes of these different taxonomic levels suggest a pattern of variation. Without the support of overt quantitative comparisons and behind the scenes mathematical operations

such as addition and the computation of ratios, Darwin's argument for the existence of variation for natural selection would have been purely speculative. But by using and comparing actual data about variation in populations, Darwin is able to make a compelling case for variation based on the qualitative work of established floral and entomological experts.

The Struggle for Existence

Now that Darwin has established the basic premise of his argument, that variation exists in natural and domesticated populations of organisms, with quantitative support, it is necessary for him to explain how this variation accumulates in natural populations. In order to accomplish this goal, he relies on arithmetical operations and quantitative evidence to support his case by rendering the difficult-to-observe action of natural selection visible for the reader. In addition, he also relies on Thomas Malthus' arguments in his *Essay on the Principle of Population* as a source of inspiration and support for the existence of competition in populations, an argument which also relies on quantitative comparisons to make its case.

For Darwin, a struggle between organisms had to exist in order for natural selection to operate. Without competition, there would be no pressure driving the spread or the extinction of variations in organic populations. He makes the importance of this condition known at the very outset of the chapter where he admonishes his reader to keep in mind the value of the struggle for existence in understanding the workings of nature.

Nothing is easier than to admit in words the truth of the universal struggle for life, or more difficult... than constantly to bear this conclusion in mind. Yet unless it be thoroughly ingrained in the mind, I am convinced that the whole economy of nature, with every fact on distribution, rarity, abundance, extinction, and variation, will be dimly seen or quite misunderstood. (52)

Once he has highlighted the importance of the struggle for existence to variation, he provides a definition of the term:

I should premise that I use the term 'Struggle for Existence' in a large and metaphorical sense, including dependence of one being on another, and including (which is more important) not only the life of the individual, but success in leaving progeny. Two canine animals in a time of dearth, may be truly said to struggle with each other which shall get food and live. But a plant on the edge of a desert is said to struggle for life against the drought.... A plant which annually produces a thousand seeds, of which on an average only one comes to maturity, may be more truly said to struggle with the plants of the same and other kinds which already clothe the ground.... In these several senses, which pass into each other, I use for convenience's sake the general term of struggle for existence. (53)

Here, Darwin explains that the struggle for existence includes not only the efforts of an individual organism to stay alive, but also its struggle to leave as many progeny as possible. In addition, he makes clear that individual survival and the survival of offspring

are influenced by competition between organisms in the same population, competition between organisms in different populations, and the struggle of an organism given the conditions of its physical environment. By providing this detailed definition with examples, Darwin reveals the complexity of the relationships that are in play in the struggle for existence.

Before Darwin can commence illustrating these different aspects of the struggle, however, he decides to provide the appropriate explanation of how such a struggle could develop in nature in the first place. For an explanation, he turns to Thomas Malthus' *Essay on the Principle of Population*:¹⁴

A struggle for existence inevitably follows from the high rate at which organic beings tend to increase. Every being... must suffer some destruction during some period of its life... otherwise, on the principle of geometric increase, its numbers would become so inordinately great that no country could support the product.
(53-54)

With these words Darwin summarizes the point made by Malthus which is that the growth of populations will eventually outstrip the amount of resources necessary to sustain them (Malthus 14). This mathematically informed argument itself is not novel to Malthus; however, he is the first, according to Donald Winch in his introduction of the *Essay*, to impart, “an added air of drama and precision to the subject by contrasting the geometric rate at which population was capable of increasing with the arithmetic rate at

¹⁴ Darwin's work drew on the 6th and final edition of the text which was published in 1826. Malthus' text was first published anonymously in 1798 with other editions appearing in 1803, 1806, 1807, and 1817.

which subsistence could be expanded” (Malthus, *An Essay on the Principle of Population* [1992], xi)

In book one chapter one of the *Essay*, Malthus makes the case that if left unchecked a population will increase geometrically using American population statistics. He chooses America as a model for population increase because at the time it was a land with boundless resources and few inhabitants.

In the northern states of America, where the means of subsistence have been more ample, the manners of the people more pure, and the checks to early marriages fewer, than in any of the modern states of Europe, the population was found to double itself for above a century and a half successively, in less than twenty-five years. (Malthus, *An Essay on the Principle of Population* [1890], 4)

In addition to using American population statistics, Malthus also turns to Leonard Euler’s (1707-83) calculations of mortality and births in which Euler also makes the case for a geometric expansion of the population.

According to a table of Euler, calculated on a mortality of 1 to 36, if the births be to the deaths in the proportion of 3 to 1, the period of doubling will be only 12 years and 4/5ths. And this proportion is not only a possible supposition, but has actually occurred for short periods in more place than one. (4)

Euler's calculations make a strong case that geometrical increase can occur in a short period of time; however, Malthus, in order to make his case seem reasonable, adopts the rate of geometric expansion based on the actual American statistics rather than the more hyperbolic rate of expansion suggested by Euler. Thus, Malthus writes, "It may safely be pronounced, therefore, that population, when unchecked, goes on doubling itself every Twenty-five years or increases in a geometric ratio" (4).

Once he has made the case for the geometric increase in population using a conservative estimate based on American population data, he then turns to his home country of England as a model for making his point that food production increases only arithmetically. He begins this portion of his argument by stating that there is no available means of determining the rate of increase in food production: "The rate according to which the productions of the earth may be supposed to increase, it will not be so easy to determine" (4).

Without available quantitative data to support his assertion, Malthus supplies his reader with a qualitative example of food production in England. He argues that in British agriculture all of the best land was already in use so that any expansion of agriculture would be into land that yields less and is more difficult to farm. As a result, "the additions that could yearly be made to the former average produce must be gradually and regularly diminishing" (6). For the sake of argument, however, he allows for the supposition that the gains to the yearly average production by the addition of these newly acquired farming acres would remain consistent. Based on this assumption, he concludes that, "the produce of this island [England] might be increased every twenty-five years by a quantity equal to what it at present produces" (6). In other words, it could only add

every twenty five years the amount which it currently produces. As a result, the best possible scenario for food production, “considering the present average state of the earth, the means of subsistence, under circumstances the most favorable to human industry, could not possibly be made to increase faster than in an arithmetic ratio” (6). Malthus forecasts the differing rates of increase as follows:

Let us call the population of this island eleven million; and suppose the present produce equal to the easy support of that number. In the first twenty-five years the population would be twenty-two million, and food being also doubled, the means of sustenance would be equal to this increase. In the next twenty-five years the population would be forty-four million, and the means of subsistence only equal to the support of thirty three million. (6)

Malthus thus employs both quantitative and qualitative assumptions and evidence to make the case that population growth if unchecked would quickly outstrip the food resources needed to sustain it. Darwin begins his arguments for natural selection based on this quantitative/qualitative reasoning. He argues that if we assume that a natural population will eventually outstrip the resources that are available for its members, then we must submit that eventually resources will diminish and a competition will ensue. Thus, a struggle for existence will develop.

Following his postulation of how a struggle for existence develops, Darwin provides his readers with quantitative evidence that in many natural populations the limit

of resources has already long been reached, and the ensuing struggle for existence has begun to operate.

Linnaeus has calculated that if an annual plant produced only two seeds... and their seedlings next year produced two, and so on, then in twenty years there would be a million plants. The elephant is reckoned the slowest breeder of all known animals, and I have taken some pains to estimate its probable minimum rate of natural increase: it will be under the mark to assume that it breeds when thirty years old, and goes on breeding till ninety years old, bringing fourth three pair of young in this interval; if this be so at the end of the fifth century there would be alive fifteen million elephants, descended from the first pair. (54)

Based on simple calculations of exponential increase in these two hypothetical situations, Darwin establishes that without the destruction of organisms or a diminution of the number of offspring that survive, populations of plants and elephants would far exceed the current numbers we observe in nature. The fact that they do not, however, is evidence that in some populations of organisms the limit between their numbers and the resources they need to sustain themselves has already been reached and the struggle for existence has taken hold.

In other cases, however, Darwin describes situations in which organisms are introduced into a new environment with less competition for resources. The result is an explosion of the organic population. This too provides evidence for the veracity of

Malthus' principle of population by suggesting that in cases where resources are plentiful populations will increase geometrically.

So it is with plants: cases could be given of introduced plants which have become common throughout whole islands in a period of less than ten years. Several of the plants, such as the cardoon and a tall thistle, now most numerous over the wide plains of La Plata, clothing square leagues of surface almost to the exclusion of all other plants.... In such cases... no one supposes that the fertility of these animals or plants has been suddenly and temporarily increased in any sensible degree. The obvious explanation is that the conditions of life have been very favorable. (54-5)

Once Darwin has provided compelling evidence that Malthus' principle of population is an established axiom of nature which accounts for the development of a struggle for existence, he then proceeds to provide his readers with detailed examples of the different manifestations of the struggle for existence. He accomplishes this by collecting and describing quantified evidence drawn from his own observations or experiments. In order to illustrate the destruction of organisms through predation and the struggle for resources, he presents the results of his observation of and experimentation on a plot of weeds in his back yard.

On a piece of ground three feet long and two wide, dug and cleared, and where there could be no chocking from other plants, I marked all the seedlings of our

native weeds as they came up, and out of the 357 no less than 295 were destroyed, chiefly by slugs and insects. If turf which has long been mown... be let to grow, the more vigorous plants gradually kill the less vigorous, though fully grown, plants: thus out of twenty species growing on a little plot of turf (three feet by four) nine species perished from the other species being allowed to grow up freely. (57)

In this description of his experiment, Darwin begins by presenting his readers with a quantitative analysis of the destruction of plants by predators that takes place during the period of observation. His numbers reveal the magnitude of the destruction. In addition, the quantification of this commonly observed relationship between prey and predator sets the stage for the more-difficult-to-observe struggles between organisms competing for limited resources. In the second portion of the experiment, he presents evidence for the operation of these more intangible processes. In order to make this invisible competition visible, Darwin tallies the number of species that appear as seedlings and then allows them to grow up. Afterwards, he calculates the number of species which survive. This strategy allows him to provide quantitative evidence that, despite the difficult-to-observe nature of this competition, a struggle has indeed been taking place between the different species of weeds for the limited resources of this small plot of land.

In the portions of the chapter which follow his turf experiment, Darwin moves his reader from simple, small-scale examples of the struggle for existence resulting from predation and the competition for resources to more sophisticated regional scenarios

involving complex natural forces. In the next part of his discussion in chapter three, he opens, as he did in the previous section, with a tangible familiar example of how changes in climatic conditions can effect the survival of populations of organisms.

Climate plays an important part in determining the average numbers of a species, and periodical seasons of extreme cold or drought, I believe to be the most effective of all checks. I estimated that the winter of 1854-55 destroyed four-fifths of the birds in my own grounds; and this is a tremendous destruction, when we remember that ten per cent is an extraordinarily severe mortality from epidemics with man. (57-58)

Though his discussion of the destruction of birds by the cold need not necessarily employ quantitative data to reveal the reality of death from extreme cold in animals, Darwin's use of a specific quantified example does have argumentative value. First, it provides his reader with a sense of the magnitude of the effect that climate can have on a species' destruction. By giving an estimate of four-fifths of the birds on his grounds and by comparing this to quantified descriptions of the number of humans that die in a severe epidemic, ten percent, Darwin creates a contrast which he hopes will connect with his readers. Second, it makes the death seem real not only by connecting it to human loss, which they can understand, but also by providing a quantified description of an actual population of birds in an actual place during an actual time, thus lending verity to evidence which might without specification seem more hypothetical than actual.

Once Darwin has provided concrete quantitative evidence to support the magnitude and reality of the direct effect that climate can have on populations of organisms, he then proceeds to give a more nuanced qualitative explanation connecting the destruction of populations through environmental factors to the struggle for existence in populations of organisms.

The action of climate seems at first sight to be quite independent of the struggle for existence; but in so far as climate chiefly acts in reducing food, it brings on the most severe struggle between the individuals, whether of the same or of distinct species, which subsist on the same kind of food.... Each species... is constantly suffering enormous destruction... from enemies or competitors for the same place and food; and if these enemies or competitors be in the least degree favored by any slight change in climate, they will increase in numbers. (58)

Here, Darwin provides his reader with a presentation of some of the abstract connections between the conditions of the physical environment and their effect on the survival of organisms. For Darwin, the effect of climatic change in nature is more than simply a condition which eliminates portions of organic populations. Instead, it is a contributing factor to the struggle for existence because it adversely affects the amount of resources available to the surviving populations, requiring the survivors to enter into an even fiercer competition for sustenance. This explanation provides his readers with a hint of the complexity in the relationships between different elements in an organic environment and their effects on the magnitude of the struggle for survival.

In the final section of the third chapter, Darwin attempts to verify the connectedness of organisms to one another and to their environment in regard to their struggle for existence. He accomplishes this by providing quantitative evidence from his own natural observations of how small changes in an “ecosystem” can have far reaching effects on the organisms that abide in them. One of the examples he gives is the introduction of a single type of fir tree to a portion of barren heath.

There was a large and extremely barren heath, which had never been touched by the hand of man; but several hundred acres of exactly the same nature had been enclosed twenty-five years previously and planted with scotch fir. The change in the native vegetation of the planted part of the heath was most remarkable... not only the proportional number of heath plants were wholly changed, but twelve species of plants... flourished in the plantations, which were not to be seen on the heath. The effect on the insects must have been still greater, for six insectivorous birds were very common in the plantations, which were not to be seen on the heath... here we see how potent has been the effect of the introduction of a single tree (59-60)

In this example, Darwin supplies his reader with quantified evidence to illuminate the nearly invisible strands that link the survival of species together. His observation illustrates to his readers the wide-reaching effects that even small changes in an ecosystem, such as the introduction of the scotch fir and the enclosures which protect

them, can have on the success of many different types of organisms in their struggle for existence.

Natural selection: calculating diversity

With the evidence and arguments in place supporting Darwin's position that diversity in nature is the result of the spread of variations through populations of organisms and that a struggle for existence takes place in nature, he can finally describe in full detail species formation by natural selection. In the fourth chapter of *The Origin of Species*, mathematical computations and representations help him to discover and provide evidence to support a new line of argument: that the more diversified an organism is, the better that organism will do in its struggle for existence.

Darwin argues that when a population reaches an equilibrium between its numbers and resources the only way that that population can continue to grow is if it is able to expand into a new ecological niche with underexploited resources. He uses the example of a carnivorous quadruped to make his point.

Take the case of a carnivorous quadruped, of which the number that can be supported in any country has long ago arrived at its full average. If its natural powers of increase be allowed to act, it can succeed in increasing... only by its varying descendants seizing on places at the present occupied by other animals: some of them, for instance, being able to feed on new kinds of prey... some inhabiting new stations. (93)

This expansion, however, requires more than just a subtle variation of character. It requires a substantial change. As a result, Darwin reasons that the same struggle for existence which influences organisms to diversify also pressures them eventually to diversify to a great degree.

Darwin's *principle of divergence of character* is an important component in describing the development of variation in nature because it answers a question which he and his critics considered a major obstacle to any theory of variation: "How do the small differences that are observable between populations of closely related species and varieties grow into the large differences that we see between genera, families, etc.?" (Browne 74). Initially, Darwin thought that variation was always minute; however, with the introduction of new mathematical techniques for calculating the boundary between large and small varieties, species, and genera he concluded that this was not always the case.

In one of the few scholarly works confirming the value of mathematics in developing Darwin's ideas, "Darwin's Botanical Arithmetic and the 'Principle of Divergence,' 1854-1858," Janet Browne argues that Darwin's quantitative analysis of populations in the second chapter also encourages him to discover the principle of divergence of character that solves the problem of increasing variation.

In the initial stages of his calculations of the ratios of varieties to species, Darwin divided the total number of organisms he was investigating into "large" and "small" groups and calculated the average number of species in the genera belonging to each of these groups. He then used the average number of species calculated for a large or a small

group as a benchmark for determining whether genera could be considered large or small (Browne 79).

Though these calculations revealed that those genera in the “large” category had the most species and that those species had the most varieties, which supported his position on the relationship between the size of a population and the number of favorable variations, these computations were based on assumptions about what constituted a large genus or species to set the benchmark for category membership rather than on calculations of the actual average number of varieties for the species and species for the genera being examined. As a result, the categories represented only assumptions of relative estimates of size rather than precise estimates based on a standardized method of comparison (79).

A communication from (Sir) John Lubbock, the son of Darwin’s neighbor at Down in the early summer of 1857, appraised Darwin that building his case on assumed estimates of size created problems in establishing rationally defensible definitions of the categories “small” and “large” for his genera. Although Darwin was initially skeptical about Lubbock’s proposed approach, he eventually did rework his estimates according to Lubbock’s suggestion with good results.

I have divided the New Zealand Flora as you suggested, there are 329¹⁵ species in genera of 4 and upwards, and 323 in genera of 3 and less. The 339 species have 51 species presenting one or more varieties. The 323 species have only 37.

Proportionately (339 : 323 :: 51 : 48.5) they ought to have had 48 1/2 species

¹⁵ This figure should be “339” but Darwin erroneously writes “329” in his letter. I am assuming that the former is the correct figure because it appears twice as the value for the number of species in genera of four and upwards.

presenting vars. So that the case goes as I want it, but not strong enough, without it be general, for me to have much confidence in. I am quite convinced yours is the right way; I had thought of it, but should never have done it had it not been for my most fortunate. (Darwin, *The Life and Letters of Charles Darwin*, 461-62)

Lubbock's additional suggestion, to work out the projected ratios of species and varieties in the larger and smaller genera and compare them to the actual ratios, provided Darwin, according to Browne, with the evidence which inspired his principle of divergence (82-83). What Darwin discovered in working out the projected ratios of varieties to species in large and small genera is that in the case of small genera there are fewer varieties than are expected or in the case of larger genera there are more varieties than expected. The result of these calculations caused Darwin to shift his focus from the quantitative relationship between the size of populations and the number of variations they produced to the qualitative nature of their variations.

This change in emphasis made Darwin shift his gaze to focus on the *success* which large genera so evidently enjoyed. He suddenly saw that it was not just variation and the fortuitous production of "good" adaptations which induced large genera to produce yet more and more species, but it was also their potency.... They [large genera] were, in fact, the very acme of success, being more widespread and more abundant in individuals than their smaller confreres, and also turning out more varieties in which "good" adaptations were likely to emerge. (Browne 86)

The realization, that larger genera not only have a greater number of species and varieties, but that this greater number of species and varieties increases the likelihood that a good adaptation will occur, led Darwin to the conclusion that a population's ability to generate diverse offspring was the key factor to their success. Thus, the mathematical calculations that Darwin undertook to determine the existence of variation in nature also led him to the realization that there was a direct correlation between the success of large genera and species and number of variations they produced.

In a letter to Hooker in August of 1857, Darwin's belief in the importance of these calculations to the development of his theory is evident.

I intend dividing the varieties into two classes, as Asa Gray and Henslow give the materials, and, further, A. Gray and H.C. Watson have marked for me the forms, which they consider real species, but yet are very close to others; and it will be curious to compare results. If it will all hold good it is very important for me; for it explains, as I think, all classification, *i.e.* the quasi-branching and sub-branching of forms, as if from one root, big genera increasing and splitting up, etc., as you will perceive. But then comes in, also, what I call a principle of divergence, which I think I can explain. (Darwin and Seward 109).

In addition to directing Darwin's attention to the importance of the potential quality of diversification, mathematics also plays an important part in providing concrete

evidence for the magnitude of the divergence that takes place in nature. As in previous chapters, Darwin chooses literally to look in his own backyard to find the data he needs.

The truth of this principle... is seen under many natural circumstances.... For instance, I found that a piece of turf, three feet by four in size, which had been exposed over many years to exactly the same conditions, supported twenty species of plants, and these belong to eighteen genera and to eight orders, which shows how much these plants differ from each other. (94)

Here, Darwin's description of a piece of turf provides quantitative evidence that different niches exist in the polity of nature and that for an organism to be successful in its struggle for existence, it must find a way to diversify so that it might exploit what is now called a different ecological niche. The staggering number of different species, genera, and orders found in a three-foot-by-four-foot piece of turf serves as a striking illustration of the number of different ecological niches that can exist even in a small area and the amount of divergence that can occur as a result of competition over resources.

Although this small-scale example is able to give his readers an idea of the magnitude of diversity among closely competing organisms, it does not provide evidence of the benefits of diversity. Darwin attempts to rectify this shortfall in his previous example by providing further quantitative evidence which supports his conclusion that a high degree of divergence from competition aids an organism in its struggle for survival. In order to make this point, he relies on a statistical analysis of the success of non-native species introduced into the United States.

In the latest edition of Dr Asa Grey's *Manual of the Flora of the Northern United States*, 260 naturalized plants are enumerated, and these belong to 162 genera. We thus see that these naturalized plants are of a highly diversified nature. They differ, moreover, to a large extent from the indigenes, for out of the 162 genera, no less than 100 genera are not there indigenous, and thus a large proportional addition is made to the genera of these states. (95)

In order to argue that variation can lead to success in competition, Darwin identifies a sample of organisms that represent highly diverse populations that have been successful at colonizing a new geographic area. In the sample he uses, roughly 62 percent of the populations identified do not even share similarities at the level of genera with one another. Once he has established that his sample is extremely diverse, he provides evidence that this diversity has led to the success of these populations. He accomplishes this with his calculation that out of the total of 162 genera represented in non-native species no less than 100 genera are exclusive to introduced species while only 62 genera have counterparts in the natural population. The large percentage of exclusively non-native genera thriving in North America, around 61 percent, provides evidence that these foreigners had more success in their adopted ecosystem because of their diversity.

In the fourth chapter of *The Origin of Species*, mathematics supplies Darwin with both inspiration and evidence to support the development of his theories of the natural world. By working out the ratios for the number of species that might be expected in a large-sized genera and the number of varieties that might be expected for a given species,

Darwin discovers that the actual numbers are higher than the ratios. This leads him to adopt the line of argument that not only do larger genera and species produce more varieties, but also that the quality of the varieties they are producing allows them to thrive more effectively by opening up new biological niches. As a result of these calculations, Darwin turns his attention to the role of diversity in the struggle for existence which encourages him to formulate the principle of divergence of character. Finally, he relies on a statistical comparison of the number of native and non-native plants thriving in North America to support this principle. The results of this comparison suggest that diversity does play an important role in the struggle for existence.

Conclusion

By the time Darwin published *The Origin of Species*, mathematical arguments and evidence had already begun to make headway into sciences outside of physics, astronomy, and geodesy. The desire for rigor and precision in all areas of natural knowledge, which had encouraged this trend, is also evident in Darwin's arguments. On inspection, these quantitative/mathematical arguments not only provide support for his position, but also inspiration. In the four opening chapters of *The Origin* where Darwin makes his case for the fundamental principles driving and directing variation in the natural world, he relies on calculations and quantitative comparisons to establish that domestic and wild organisms can undergo variation, that a fierce struggle for existence exists, and that organisms whose variation is more pronounced than others tend to be more successful in the struggle for existence. In all of these cases, the simple operations

of quantitative comparison as well as addition, subtraction, division, multiplication, and ratios suffice to offer illumination of and persuasive arguments for the nearly invisible forces of struggle, variation, and diversity that slowly shape and reshape the façade of nature.

Although the mathematical elements of the arguments in *The Origin* may pass undetected by modern scholars accustomed to the sophisticated mathematical models of twentieth century biology, under the scrutiny of his nineteenth-century contemporaries such as Francis Galton (chapter four) and Karl Pearson (chapter five), Darwin's arguments supplied inspiration to develop novel and more sophisticated mathematical approaches to the study of variation, evolution, and heredity. By looking at his work in this larger context, it is possible to see it as an encouragement, if not a precedent, for the use of mathematically informed arguments in the study of these biological phenomena.

Chapter 2: Drawing Back the Veil: Probability and Statistics from 1600 to 1900

*It is of primary importance to keep out of view man as he exists in an insulated, separate, or in an individual state, and to regard him only as a fraction of the species. In thus setting aside his individual nature, we get quit of all which is accidental, and the individual peculiarities, which exercise scarcely any influence over the mass, become effaced of their own accord, allowing the observer to seize the general results. –Adolphe Quetelet in *A Treatise on Man and the Development of his Faculties*.¹⁶*

The evidence provided in the previous chapter suggests that Darwin used quantitative comparisons of data describing organic populations to make his arguments for variation, natural selection, and the principle of divergence of character in *The Origin of the Species*. In chapters four and five, I will show how this method of argument influenced later researchers such as Francis Galton and Karl Pearson to look for inspiration and support for their hereditary theories in the special topics¹⁷ of probability and statistics. Before examining the role of these special topics in their arguments, however, a preview of the general concepts and methods and the socio-historical context in which they develop is in order. This chapter supplies a brief introduction to the basic philosophy and methodology of probability from its birth in the middle of the seventeenth century to the period right before its transformation at the beginning of the nineteenth century.

In addition, this chapter offers a more detailed account of conditions in the opening half of the nineteenth century which expand the philosophies and methods of probability from astronomy and geodesy to human physiology, psychology, and society.

¹⁶ Adolphe Quetelet, *A Treatise on Man and the Development of his Faculties* (Gainesville, Florida: Scholar's Facsimiles & Reprints, 1969) 5.

¹⁷ Special topics are lines of arguments developed for use in a specialized field, or as Aristotle states in *The Rhetoric*, “there are also those special lines of argument which are based on such propositions as apply only to particular groups or classes of things” (Aristotle 31).

This expansion represents a revolution in probability and statistics and creates the context for the application of probability to the study of evolution, variation, and heredity and the debates leading to the establishment of a modern genetic paradigm informed by the methods and philosophies of probability.

The Seventeenth Century: Pascal, Fermat, and Huygens

The earliest period in which probability theory was coherently developed, known as the period of classical probability, began in the middle of the seventeenth century and extended into the middle of the eighteenth century. In *Classical Probability in the Enlightenment*, Lorainne Daston characterizes probability in this initial stage as a rational calculus developed to combat the skeptical philosophy that absolute certainty was beyond human grasp (Daston xi). This new “calculus of good sense” drew inspiration from legal situations in which decisions concerning equity had to be made based on unknown future conditions (14). Gambling problems, for instance, served as the source of discussion in the texts in which an organized concept of probability theory originated. The letters sent between Blaise Pascal (1623-62) and Pierre de Fermat (1601-65) in 1654 regard a question of stakes in an interrupted game of chance put to Pascal by the Chevalier de Mere. In these letters, Pascal and Fermat discuss the hypothetical situation presented by De Mere where a three point game¹⁸ is being played in which each player wagers 32 pistoles. The game is interrupted when player *A* has won two games and player *B* one. The interruption raises the question, “How should the stakes be divided?” Pascal argues

¹⁸ A three point game is essentially a series of three or more hands of cards in which the first player to win three hands wins the pot.

that since player *A* has won two games already even if he were to lose in the next round he would still have walked away with half the pot, 32 pistoles. This leaves 32 pistoles up for grabs. Pascal reasons that because each man has a fifty percent chance of winning or loosing the remaining 32 pistoles, this portion of the pot should be divided in half. Consequently, player *A* should get 48 pistoles (32+16) and player *B* should receive 16 (Daston 15-16).

In the exchange between Fermat and Pascal, two central concepts of classical probability, *equality of chance* and *expectation*, play a central role in reasoning about probable outcomes. Expectation is evident in Pascal's assumption that player *A* should automatically receive half of the pot because whether he won or lost in the next hand, he would still be entitled to half. Equality of chance is also present in the assumption that either player has an equal chance of winning or losing in the next round. This is the rationale behind his argument that the remaining half of the pot should be split between the two players.

At around the same time of Fermat and Pascal's correspondence, Christian Huygens (1629-95) was developing his own mathematical treatments of gambling puzzles which were published in his 1657 treatise *De Ratiociniis in Aleae Ludo* (Computations in Games of Chance). In this first published treatise on probability, Huygens, like Pascal and Fermat, also creates a calculus of expectations which deals with, "the division of stakes or the 'reasonable' price for a player's place in an ongoing game" (24).

The central axiom around which Huygens develops his theory of expectation is that of the fair game: "I take as fundamental for such games that the chance to gain

something is worth so much that, if one had it, one could again get the same chance in a fair game, that is, a game in which nobody stands to lose” (quoted in Hald 69). According to Anders Hald in *History of Probability and Statistics*, Huygens’ explains the axiom through example:

Suppose that somebody has 3 shillings in his one hand and 7 in the other and that I am asked to choose between them; this is so much worth to me as if I had 5 shillings for certain. Because if I have 5 shillings I can establish a fair game in which I have an even chance of getting 3 or 7 shillings, as will be shown below.
(69)

In this passage, Huygens reasons that with five shillings he can establish a fair game if the possible outcomes are to receive three or seven shillings in exchange for staking five shillings in the playing of the game. This is a fair game because 1) the player begins with an equal chance of receiving either the three or the seven shillings and 2) the possible gain from the bet is equal to the loss (both are two shillings). Though this explanation is slightly different from the rationale discussed by Fermat and Pascal, the essential parameters are the same: there is an expectation of a fair balance between risk and reward and either player has an equal chance of winning or losing.

Most, if not all, of the early discussions of probability, like those of Pascal, Fermat, and Huygens, were limited to problems of gambling (Chatterjee 157). The type of probability dealt with was strictly *a priori probability*, meaning that the calculations of expectations were made only in cases where the number of possible outcomes was known

before the calculations of probability were made. (In Huygens', Pascal's, and Fermat's examples there are two known outcomes. In Huygens' game, three or seven shillings can be won; and in Pascal and Fermat's game, there is assumed to be a winner and a loser) (Daston 26). The outcomes in any trial are considered *equiprobable*. In other words, it was just as likely that one outcome would happen as the other. This assumption is at the heart of the concept of a "fair game," which Huygens employs as the fundamental axiom for his theory of probability.

Unlike the modern sense of the term "chance," which conveys a sense of randomness, the term "chance" in these early treatises had nothing to do with events which occurred without rhyme or reason. Instead, the concept was used to describe a preordained event whose precise outcome was unknown but was still in keeping with a predetermined set of outcomes (10).

Jakob Bernoulli and Abraham de Moivre

In classical probability, the mathematics was limited primarily to problems of gambling that involved very simple situations in which the degree of probability of an outcome was assumed known: for example, it was assumed that one player at a game of chance was as likely to win as another. In the late seventeenth and early eighteenth century, however, probabilists and philosophers of logic began to complicate probability by introducing models in which the probability of an event could not be assumed.

The problem of assumed or a priori probability was raised initially by Pascal in what is understood to be his work in the Port Royal *Logic*¹⁹ (1662) and by the German mathematician/philosopher Gottfried Leibniz (1646-1716) in his philosophical discussions of jurisprudence. Essentially, these works explore situations in which rational decisions must be made with insufficient evidence. The decision maker begins with two choices: guilty or not guilty in Leibniz's case and to lead a pious life or to live a worldly life in Pascal's. In these two choices, there exists a priori a true and a false answer. The problem faced by the decision maker is that though there is a correct answer there is an insufficient amount of evidence to identify which of the two options is correct. In Leibniz's case, there is insufficient evidence to say with certainty whether the accused is guilty or not, and, in Pascal's case, there is an insufficient amount of evidence for the existence of heaven or hell. Because the true a priori probability of each event is not known, either 100% certain or 100% uncertain, the decision maker must decide based on the knowledge they have at the moment a decision has to be made. In other words, they must develop a subjective probability based on their own experience (Chatterjee 164-67).

Though Pascal and Leibniz were the first to introduce the complication of insufficient evidence into the assessment of probability, the first person to develop mathematical tools to deal with this problem was Jakob Bernoulli (1654-1705). In his book *Ars Conjectandi (The Art of Conjecture)*, published posthumously by his nephew Nicolas Bernoulli in 1713, Jakob Bernoulli developed a system for dealing with probability which operated on the assumption 1) that we can reach a degree of reasonable certainty about the a priori probability of an event and 2) that this can be accomplished

¹⁹ The title of this text is actually *Ars Cogitandi (The Art of Thinking)*, however, the Port Royal *Logic* is the more popular title adopted for the text.

through the use of many repeated trials (Chatterjee 168). In chapter four of *Ars Conjectandi*, Bernoulli describes these principles mathematically in what is known today as “Bernoulli’s limit theorem.” The theorem makes the case that the calculated a posteriori probability of an event p gets closer to the true a priori probability of an event P the greater the number of trials n that are conducted.

In his text, Bernoulli uses the example of a sealed urn of balls with a fixed ratio of black to white balls from which drawings are made at random and with replacement. Each ball is as likely to be chosen as any other (Daston 238). He argues that if the ratio of black to white balls is *known a priori*, we can calculate the number of trials necessary to, “ensure that the ratio of black-to-white balls falls within a certain margin of the true ratio with a given probability” (238). The limit theorem then provides the means of calculating the number of trials it would take to reach a specified degree of certainty about the underlying predetermined ratio of black-to-white balls.

Although Bernoulli’s theorem allows that a high degree of certainty can be established through a large number of trials, he recognizes that it is never possible to reach complete certainty about an event, moving probability from the purview of the dialectic to the jurisdiction of the rhetorical. He suggests instead that a “moral certainty” can be established through probabilistic analysis.

That is *morally certain* whose probability nearly equals the whole certainty, so that a morally certain event cannot be perceived not to happen; on the other hand, that is *morally impossible* which has nearly as much probability as renders the certainty of failure moral certainty. Thus one thing is considered morally certain

which has 999/1000 certainty, another thing will be morally impossible which has only 1/1000 certainty. (Hald 248)

Bernoulli's limit theorem reflects standard seventeenth- and eighteenth-century beliefs about nature. Like his predecessors, Bernoulli assumes that a certain order exists in nature. This belief is evident in his limit theorem which relies on the existence of a fixed and quantifiable a priori ratio of outcomes, even when the ratio is unknown, in order to calculate the relationship between the number of the trials and the degree of certainty that the outcome of those trials will point to the correct ratio.

Bernoulli's belief in the transcendent order of things and his mathematical solution to finding patterns in nature through collective independent trials offered a method for natural philosophers seeking to decode the book of nature through empirical induction, and, according to Shoutir Chatterjee in *Statistical Thought*, a rationale for the development of a program of natural investigation joining statistics,²⁰ the collection and presentation of quantified information, and probability, the calculation, using the mathematical principles of probability, of the degree of certitude for an event based on a collection of quantified outcomes (Chatterjee 171).

In addition to supporting induction, Bernoulli's limit theorem inspired the further development of both statistics and probability in the first half of the eighteenth century. Some of the most notable contributions based on his theory come from Abraham de Moivre (1667-1754), who is credited with developing the concept of a normal

²⁰ In this definition I am not using the modern sense of the term statistics which describes a discipline in which both probabilistic assessment and statistical collection are united.

distribution and with creating a formula by which the probability p of an event given a large number of trials n can be calculated (Daston 250).

In *Miscellanea Analytica* (1730) and *The Doctrine of Chances* (1718, 1738, and 1756), De Moivre solves the problem of calculating a posteriori probability for a large number of trials. His solution involves, “finding the probability... when n symmetric dice with f faces marked $1, 2, \dots, f$ are thrown” (Chatterjee 176). The function by which he solves this problem comes to be known as a probability generating function. Creating a function by which the probability of an increasingly large number of trials can be calculated was an important contribution to the study of probability because, previous to its introduction, calculation of probabilities for a large number of trials was extremely arduous (Gonik and Smith 79). With this new relatively simple technique in hand, it was, “practicable to apply probability theory to an indefinitely large number of independent events” (Porter 93). This provided for both a practical and theoretical expansion of probability to work with large statistical samples and to create hypotheses about the number of trials required for high degrees of accuracy in induction.

In addition to solving the problem of computing the probability in cases with a large number of trials, De Moivre is perhaps more importantly credited with conceiving of the normal distribution, though it is clear that he himself did not fully understand the significance of his contribution. By deploying the newly invented tools of calculus, “De Moivre showed that when p , the assumed probability of an event was equal to $.5$, the binomial distribution was closely approximated by a continuous density function which could be described very simply” (Gonik and Smith 79). This density function, called in

modern parlance the “central limit theorem,” describes a distribution which is bell shaped with a mean of 0 and an area under the curve of 1.

The early eighteenth-century treatises of Bernoulli and De Moivre represent the first important forays into the development of inductive probability. Bernoulli’s limit theorem established quantitatively the degree of reasonableness of an inductive conclusion reached through repeated trials. This ability to quantify the degree of certainty for induction added a degree of validity to inductive reasoning. In addition, De Moivre refined Bernoulli’s work adding a methodology for calculating the probabilities for a large number of trials. This methodology expanded the application of Bernoulli’s limit theorem and made it more practical to apply in assessing experimental outcomes. Also, De Moivre was the first to describe the bell shaped binomial distribution which was to take center stage in the nineteenth century as the “normal” distribution for a large range of natural phenomena.

Pierre-Simone Laplace

The probabilistic endeavors of the seventeenth and early eighteenth centuries to reach a degree of moral certainty about the a priori probability of a phenomenon through repeated trials continued to influence the study of induction in the late eighteenth and early nineteenth centuries. Importantly, it influenced the work of Pierre-Simone Laplace, who tried to develop a mathematical means of identifying the shape of the distribution of error and the limits for a homogeneous distribution of causes as well as means for identifying true value for a series of observations.

Between 1799 and 1825 Laplace published his five volume *Treatise on Celestial Mechanics*²¹ (*Traite de Mecanique Celeste*) whose primary goal was, “to show how all celestial phenomena derive from the principle of universal gravitation” (Hahn 144). In book three volume one, Laplace addresses the problem of how to assess whether or not a comet’s orbit was parabolic. Laplace’s solution was to formulate an orbit shape from the known data points with the smallest range of error. If the shape was parabolic, then his hypothesis stood. If it was not, then the hypothesis of a parabolic figure could not stand for the given data within the limits of observational error (Gillispie 188).

Laplace’s earliest work on statistical induction in *Celestial Mechanics* sought to determine the range of observational error on the basis of calculations of the distribution of values gathered from actual observations of comet trajectories. Finding this method intractable, he abandoned it in favor of an approach based on a priori probability. The point of view regarding the range of observational error he developed with this approach was similar to that used in statistical sampling in which the mean error in a large number of observations falls within predetermined limits (218). What he hoped to do was to determine the limits of error, its distribution, and the true value of a measurable phenomenon. In order to do this, he developed what became known as the “central limit theorem” which made the case that if observations, “are repeated an indefinite number of times, their mean result converges on a limit such that, if an equal interval on either side be made as small as one pleases, the probability that the result will be contained therein can be brought so close to certainty that the difference is less than any assignable magnitude” (Gillespie 218).

²¹ Volumes one and two were published in 1799; volume three in 1802, volume four in 1805 and volume five in 1825.

In this new approach, Laplace made the case that there is regularity in the sources of errors in astronomical observation. Assuming that the “real value” of an observation is the mean value of the curve where the probability of error approaches 0, there is an equal probability of over or underestimating the true value. Also, he suggests that because big mistakes are rarer than small mistakes in observation, it is more probable that a given measurement is closer to the true value of an observed phenomenon than farther away from it. All of these parameters are illustrated in the bell curve where the actual value is located at the highest point in the curve being the observation most frequently made, and other observational values fall in frequency in equal measure from this maximum value down both sides of the curve, hence above or below the true value (Daston 271).

Underpinning Laplace’s mathematical description of the relationship between the “real value” of observations and observational errors is the belief that even the “errors” themselves are the result of an unknown mixture of causes which have a regular pattern. Daston explains,

Laplace envisioned nature as a composite of “regular” (or “constant”) and “irregular” causes. Even the irregular causes exhibited a collective regularity, however: in the long run their effects were symmetric and canceled one another out, revealing the steady operation of the underlying constant cause. (270)

The “irregular” causes described by Daston here are synonymous with the category of “errors” in the sense that they are false outcomes which interfere with uncovering the “true value” of an observation. Drawing on Bernoulli’s work, Laplace

makes the case that the more observations that are conducted, the easier it is to separate the regular or true value of an event from the series of irregular or false outcomes.

Laplace's probabilistic model of true outcomes versus error supported a philosophy of determinism in which the physical universe was conceived of as an intricate mechanism which functioned according to predetermined laws. Because of the imperfection of human senses, however, these inner workings are largely obscure and can appear chaotic to human observers. Despite this appearance of randomness, Laplace firmly believed that every entity, from the smallest atom to the largest cosmic body, obeys predetermined laws. This belief is evident in the first section of his *Essai Philosophique* where he writes:

An intelligence which for a given instant knew all the forces by which nature is animated and the respective situations of the existences that compose it; if further that intelligence were vast enough to submit these given quantities to analysis; it would embrace in the same formula the greatest bodies in the universe and the lightest atom: nothing would be uncertain to it and future as the past would be present before its eyes. (Laplace 2)

Based on his conception of errors having irregular sources and the size of errors as a means by which to judge whether or not a researcher had discovered the true value of a phenomenon, Laplace developed a method for testing whether differences in the data gathered in separate observations to prove the same hypothesis were sufficiently significant to warrant further investigation into the possible causes of the differences. To

illustrate his method, Laplace took on what was for eighteenth-century statisticians and probabilists a popular set of issues: “the problems . . . related to sex ratios at birth” (Chatterjee 214). Using his formulae for estimating the probable outcome of a future trial based on the outcomes of past trials, he estimated that the probability of a male child being born in Paris rather than a female child was greater than 50 percent ($p > 1/2$). Then, using his estimation of the distribution of error, he found that the chance that this estimation was wrong (i.e. $p \leq 1/2$) was 1.5×10^{-42} . This led Laplace to conclude, “that ‘one can regard it as equally certain as any other moral truth’ that for Paris $p > 1/2$, or in other words, nature favors more the birth of boys than of girls” (quoted in Chatterjee 214).

Although John Arbuthnott (1667-1735) and Daniel Bernoulli (1700-82) preceded Laplace in their efforts to employ hypothesis testing in their own work, Laplace was the first to explicitly test hypotheses and to provide the mathematical tools for doing these tests. Laplace argued that these mathematical tests were an important element in a disciplined approach to induction because they allowed the natural philosopher to determine mathematically whether a difference between statistical results was significant enough to encourage further investigation. This determination helped them avoid what Laplace described as “vain speculations” about whether observed differences were significant or not (Chatterjee 215).

In his work, Laplace made significant strides in developing practical methods for testing hypotheses, assessing the distribution of errors, and determining the true value of observations. These tools, though rough, strengthened reasoning from induction because they allowed researchers to describe quantitatively the probability of their conclusions

and the validity of their programs of research. In addition, these tools were important because they provided the mathematical support for Friedrich Gauss's law of error which represents for this study an important watershed in the development of a relationship between inductive probability and the study of evolution, variation, and heredity in the nineteenth century.

Probability and Statistics in the Nineteenth Century

The methods that Laplace developed in order to more accurately assess the true values of celestial observations were also used by him and others to determine with greater precision the shape of the earth. The study of the earth's shape, called "geodesy," was an undertaking of national importance because it led to improvements in navigation which impacted military and economic concerns. The expansion of probability and statistics into geodesy in the eighteenth century was a product of its obvious economic and military value as well as its close connection to astronomical observations. In the opening of the nineteenth century, however, the applications of probability and statistics expanded to fields such as human physiology, behavior, and psychology which were less obviously beneficial than geodesy and less closely related to astronomy. This new era for statistics and probability was ushered in by 1) the increasing availability and range of statistical data, 2) the development of the "law of errors" by Friedrich Gauss (1777-1855), and 3) the vision and energy of astronomer Adolphe Quetelet (1796-1874). The expansion of probability and statistics in physiology, human behavior, and psychology as a result of these factors accounts in large part for the influence of the philosophy and

methodology of probability and statistics on investigations of evolution, variation, and heredity.

One of the most important developments in the nineteenth century affecting the expansion of the mathematics of probability is what Ian Hacking calls “an avalanche of numbers” (5). Although various religious and political institutions had been gathering statistical information about human populations since the seventeenth century, the amount and availability of that information was extremely limited.

In the opening decades of the nineteenth century, however, the statistical flood gates opened and there was a deluge of data pouring from the government printing houses of Western and Eastern Europe. Part of the reason for this sudden increase, according to Ian Hacking, was that the Industrial Revolution, in full swing at the beginning of the nineteenth century, required greater precision in measurement (Hacking 5). In order to develop a transportation infrastructure, for example, standards had to be described and released to industry for values such as the width and height of train cars and their towing capacity. In addition to the economic exigence for releasing statistical data, Hacking cites the period of peace that descended on Western Europe following the end of the Napoleonic wars as an influence on the amount of data which was made publicly available. With the end of the war there was less fear among Western European states that their own statistical data could be used against them. As a result, they felt more comfortable about releasing statistical data, making more information open to the general public for scrutiny and academic investigation (Chatterjee 267).

The avalanche of numbers precipitated by peace and trade provided the material for a probabilistic revolution; however, this revolution also required skill, innovation, and

inspiration to expand the application of probability from the study of gambling, geodesy and astronomy to investigations of biological phenomena other than mortality and birth. Arguably, two of the most influential suppliers of these ingredients were Friedrich Gauss and Adolphe Quetelet.

At the beginning of the nineteenth century, an important advancement in the mathematics of probability paved the way for its expansion into the study of evolution, variation, and heredity. This advancement was the development in 1809 of a mathematically defensible “law of error” by German mathematician-scientist Friedrich Gauss. As an astronomer, Gauss first came into contact with probability as a means by which he could manage the data from his observations of the heavens (Chatterjee 225).

For Gauss, as for Laplace, there were two kinds of errors: errors which were irregular—the result of a collection of numerous imperceptible causes that were always present—and errors which were regular or systematic—the result of inaccuracies introduced by conditions, instruments, etc. (Hall 73). With an eye towards Laplace’s work on error and hypothesis testing, Gauss believed that systematic errors could be discovered if researchers could accurately describe the distribution and boundaries of irregular errors. By comparing the limits and distribution of these errors to the distribution of their observational values, astronomers could tell whether these values were within the range of the norm for error or whether they fell outside that range and were, therefore, under the influence of some systematic or significant causal variable.

Although it was a widely held belief at the beginning of the nineteenth century that the “true value” of a set of observations was the arithmetic mean of a sample of independent measurements, there was still as yet no satisfactory way of proving that the

mean of a set of observations was in fact its true mean (Chatterjee 225). The major problem that remained to be solved was that there was no tenable index of error against which the distribution of observational results could be checked to find out whether or not the distribution was within the limits of acceptable error.

Gauss's primary contribution to probability was that he successfully calculated the limit and distribution which random errors should ideally take. His approach to the problem was somewhat different from Laplace's who had tried earlier but failed to come up with a generally applicable distribution of error. Whereas Laplace began with models of error and then tried to fit them to models of observational results, Gauss began by assuming theoretically that the mean of an observational sample is a good estimate of the "true value." Starting from this assumption, Gauss then considered what form of density function ($f(\cdot)$) coincided with the mean in the observational sample (Chatterjee 226). The answer, based on this method of calculation, was the normal distribution described earlier by De Moivre.

Gauss' development of the "law of error" represented a watershed in probability because for the first time a mathematical method existed whereby the "true value" of a set of observations could be mathematically substantiated by comparing it to a predetermined distribution of errors. This index of verity could, on the one hand, be a powerful tool for identifying fixed patterns in nature and, on the other, the true value of any measurable phenomena. These capabilities made it attractive to astronomers and geographers in their study of the heavens and earth; however, they also supported a greater potential for general philosophical interest in the possibility of demonstrating other underlying patterns in nature.

That the mathematics of probability was ripe for applications in other fields does not, of course, explain why it expanded from specific applications in geodesy and astronomy to the study of a wide range of human physiological, behavioral, and psychological phenomena. That probabilists had already conceived of applying probability to the study of human social and biological phenomena is already clearly indicated in its history. Though, at its inception, probability was linked to assessing legal questions of fairness in contracts and gambling disputes, later probabilists such as John Arbuthnott and Laplace used it to determine human birth ratios.

Though these earlier applications are directed towards some aspects of human phenomena, they do not automatically warrant the application of probability and statistics to all aspects. Whereas the outcomes of games of chance and births follow relatively stable patterns of frequency and are, therefore, amenable to probabilistic analysis, other aspects of humanity, such as the psychological dispositions of individuals and the pattern of inherited traits, are not. As a result, before probabilists would venture out into the murky worlds of human physiology, psychology, and behavior, they needed to be convinced that there was enough stability in these categories to warrant their efforts.

With the growth of available statistical data and a widening range of statistical subjects, a latent conviction began to develop that there were underlying patterns in these phenomena; however, it required someone to apply the mathematics of probability to non-traditional categories of human data and to produce persuasive evidence and arguments that regular patterns could be found in this data. That someone was Adolphe Quetelet—a Belgian astronomer who applied ideas and methodologies from probability and statistics, specifically Gauss’s law of error, to human populations for the purpose of

constructing laws describing underlying regularities in human physiology, psychology, and behavior.

During a short visit to Paris in 1823, Quetelet came under the influence of [Joseph] Fourier and Laplace whose works introduced him to the basics of probability and the theory of errors which he was inspired to use to study human populations (Chatterjee 270). He explores this application in his groundbreaking book *Sur l'homme et le developement de ses facultés, essai d'une physique sociale* (1835) (A Treatise on Man and the Development of his Faculties, an Essay on Social Physics) in which he makes the case that the theory of error can be used to describe human attributes. He assumes that despite the differences across individuals in a human population there are underlying regularities within the whole. From this assumption, he argues that for any relatively homogeneous population, or race, there is a stable relative frequency of a particular trait which represents a characteristic of that group. By combining all the average frequencies of all the quantifiable characters, it is possible to determine “the average man” (l’homme moyen) for a particular homogeneous population (271).

In a later 1844 monograph entitled *Sur l’appréciation des documents statistiques, et en particulier sur l’application des moyens* (An Appreciation of Statistical Documents in Particular on their Application of Averages), Quetelet provided evidence for his theory by studying the actual distribution of characteristics in a real human population (Hacking 108). Specifically, he analyzed the distribution of chest measurements of nearly 6,000 Scottish soldiers which had been collected by a military tailoring contractor and had been recorded in the *Edinburgh Medical Journal* in 1817 (Chatterjee 271, Hacking 109). He took the measurements and fitted them to a normal symmetric binomial distribution

(Chatterjee 271). The results confirmed a satisfactory fit between the two. This meant that the chest size of soldiers distributed according to the “law of errors” (272).

The fact that Quetelet’s measurements distributed in the shape of a normal binomial distribution described by the law of errors supported his assumption that there was an underlying regularity within homogeneous populations and that the mean of his sample represented the “true value” of Scottish men with regards to their chest size. He argued that the regular differences, i.e. errors, in the distribution of values could be accounted for by “accidental causes” such as climate, nutrition, etc. which interfere with the underlying “constant causes” of nature that act in a predictable fashion to achieve a “true value” for a particular trait (Porter 108).

Quetelet’s conceptualization of variation in the biology of humans in the same terms as errors of observation from astronomy and geodesy is a rhetorical tour de force which relies on the development of an analogy between observations of the heavens and the earth and observations of human physiology. In the 1844 monograph he begins to make his case for this analogy with the statement, “Let us suppose that I can measure the height of some individual several times, with great care” (108). He follows this with the assertions, “The measurements won’t be identical. If the causes of error work equally towards measuring high and low, there will be a distribution with values clustering around the average height. There will also be a dispersion measured by probable error” (108).

In the next step of his argument, Quetelet makes the case that this procedure for measurement is the same as the procedure for measurements made in astronomical observations: specifically to measurements made at the Greenwich Observatory. As

Hacking describes Quetelet's explanation, "There [in astronomical observation] we have mean, probable error, and the whole Gaussian analysis. This established practice is exactly analogous, he [Quetelet] said, to measuring the height of one man over and over again" (Hacking 108).

At this point, Quetelet makes an important leap from thinking of the height of an individual as a single value which can be measured many times to height as a single value within a larger series of measurements of a population. He suggests that if, in measuring many examples of a certain type of phenomenon, the overall distribution of values follows the distribution of the law of error, then we can assume that the "collection of statistics is derived from a single homogeneous population defined by a real quantity" (109). If, on the other hand, the values do not distribute in terms of this scheme, then the population being investigated is most likely made of "several distinct but mixed populations" (109).

Quetelet's argument and proof that the law of errors could be used to identify a "real quantity" for a population of humans had important philosophical and rhetorical ramifications. It allowed that particular features of human populations could be objectively described through the compilation of statistical data and the analysis of that data using the law of errors. This allowance provided support for the idea that like natural physical phenomena, human biological phenomena occurred in knowable law-like patterns. This revelation would engender later nineteenth-century programs to investigate not only the regularities in human biology, but also the regularities in human psychology and behavior. Though Quetelet himself did not pursue actual data analyses on these last two sets of phenomena, he did suppose that, "even non-measurable normal characteristics

such as propensity ('penchant') for acts like marriage, drunkenness, suicide, or crime were similarly distributed over the group [i.e. homogeneous racial population] around a true value representing the 'average man' in the group" (Chatterjee 272).

Quetelet's statistical method and his goal of using it to establish a social physics were well received in Belgium and across the channel in England. His work was read by a generation of British scientists and social philosophers who embraced his philosophy and methodology and attempted to apply it to understanding society and nature. His followers included a wide range of luminaries from historian H.T. Buckle, who authored a history of civilization which sought to establish that the course of human societies was rigidly determined by conditions of climate and locality, to physicists Clerk Maxwell (1831-1879) and Ludwig Boltzmann (1844-1906), who used Quetelet's simplification of the heterogeneous individual characteristics of a population into a statistical "average" to explain the behavior of molecules in a body of gas under heat transmission in terms of the average probabilistic speed of a molecule (278, 281).

In addition to contributing to the theoretical discussion about the usefulness of probability and statistics in examining human behavior, Quetelet also had a substantial impact on developing institutions for statistical-probabilistic research. In an 1833 meeting of the British Association for the Advancement of Science, he took a leading role in successfully advocating for the formation of a separate statistics section (275). As a result of his successful campaign, the Statistical Society of London, the progenitor of the Royal Statistical Society, was formed the following year in March of 1834 (Cullen).

Along with his instrumental role in the formation of a national statistical society in England, and a Central Statistical Commission in Belgium in 1841, Quetelet was also a

leading force behind the establishment of an international statistical congress. In 1853 the first such congress met in Belgium under his presidency to, “standardize the terminology and methods of collection of official statistics in different countries and make them internationally comparable” (275). This body was the predecessor of the International Statistic Institute, which was established in 1885.

Conclusion

In order to understand the source of supporting arguments in debates about the nature of evolution, variation, and heredity, and the intellectual and social context in which these debates developed, it is necessary to understand both something about the general history of ideas in probability and about specific developments in the fields of statistics and probability that led to the introduction of probability into the study of heredity, variation, and evolution in the nineteenth century. The general history provided in this chapter offers a glimpse into the methodological and philosophical developments in probability from which arguments regarding the nature of variation and heredity will be made in the latter half of the nineteenth century. Chapter three will show that the concepts of equiprobability and a priori probability established in the early classical period and carried on throughout the eighteenth and nineteenth centuries play a central role in the invention and delivery/development of Mendel’s arguments in “Experiments in Plant Hybridization.” Chapter four will reveal the importance of Gauss’s “law of error” and Laplace’s philosophy regarding “regular” and “irregular” errors in the development of Francis Galton’s theory of heredity. Finally, the fifth chapter will examine the use of

the law of error and Bernoulli's limit theorem by Karl Pearson to support arguments in favor of a statistical\probabilistic approach to the study of heredity and evolution. It will also include a discussion of how a more cautious assessment of the descriptive power of the normal curve is used by Pearson's adversary, William Bateson, to challenge the validity of his position. These examples provide evidence of the emergence and the spread of a new line of argument for researchers hoping to develop and support their theories of variation, evolution, and heredity as well as illustration of the rhetorical problems attached to the application of the operations, formulae, and principles of probability and statistics to making arguments about these phenomena.

Chapter 3: Hidden Value: Mendel, Mathematics, and the Case for Uniform Particulate Inheritance

In the study of evolution progress had well-nigh stopped.... Such was our state when two years ago it was suddenly discovered that an unknown man, Gregor Johann Mendel, had, alone, and unheeded, broken off from the rest—in the moment that Darwin was at work—and cut a way through. William Bateson in Mendel's Principles of Heredity.²²

Charles Darwin's *The Origin of Species* provided arguments and evidence which paved the way for a more sophisticated understanding of organic forms and their relationships with each other and their environment. Through the use of quantified evidence and arithmetical calculations, he offered arguments and evidence that organisms varied and that variation could spread through populations and accumulate to such a degree that organisms would physically diverge from their progenitors.

Despite Darwin's important work describing variation and its spread throughout populations, there were still important pieces of the puzzle missing. He had no reasonable theory for the mechanism of inheritance nor could he answer critics such as Fleeming Jenkin, who pointed out that if variations were small and needed to be accumulated regularly over vast amounts of time, they would become swamped, blended away over repeated generations of cross-breeding within the normal population (i.e. the majority of the population which does not exhibit a particular variation). Although Darwin did offer his theory of pangenesis, in which he argued that different characteristics of the parents were transferred to their offspring by gemmules in the blood, this theory was unproven and could not explain how characters could remain stable over time.

²² William Bateson, preface, *Mendel's Principles of Heredity: A Defense* (Cambridge: Cambridge UP, 1902) v.

Unknown to him, and the majority of natural philosophers at the time, the problem of character stability was being worked out by a monk laboring over carefully selected and organized samples of peas. The monk, of course, was Gregor Mendel and the carefully planned series of breeding experiments that he undertook from 1856 to 1864²³ provided the evidence for what we now recognize as the first modern theory of genetic inheritance.

Although Mendel had described in the published findings of his research “Versuche über Pflanzen-Hybriden” (“Experiments in Plant Hybridization” 1865) the basic pattern of inheritance of seven separate traits in peas, the value of his work was not recognized until his findings were simultaneously “rediscovered” by three researchers, Hugo De Vries, Carl Correns, and Erich von Tschermak, in 1900. Since the rediscovery of his findings, Mendel has earned a top place along with Darwin in the pantheon of great thinkers in the science of evolutionary biology, as both Darwin’s theory of evolution and Mendel’s theory of inheritance are necessary complements to one another in providing a full picture of the process of evolution.

From a rhetorical perspective, Mendel’s work and its reception elicit two important questions: 1) why were Mendel’s arguments so compelling to readers such as William Bateson who would, in the beginning of the twentieth century, successfully persuade the English biological community to accept Mendel’s theory of uniform particulate inheritance? And 2) why to his contemporaries was his argument so unpersuasive that only one of them ever bothered to respond to it critically and only a few even mention his work until its rediscovery in the twentieth century?

²³ These dates vary in different historical accounts from as early as 1854 to as late as 1865. I have chosen these dates based on majority opinion in the sources (Henig 2000, Iltis 1932, Olby 1966, Orel 1984).

The answer to the first question, I believe, can be found in Mendel's counterintuitive results and his unorthodox methods of arguing from the principles of probability and combinatorics.²⁴ Whereas Darwin held that species varied over time, a widely held belief among breeders, Mendel believed that species characteristics remained constant. Mendel's particular take on heredity allowed for traits to remain stable over time and to accumulate in a population whereas in Darwin's model they might quickly erode due to interbreeding. In addition, though Darwin does rely on quantitative comparisons to support and discover arguments, Mendel relies on a more robust program of argument using mathematical and quasi-mathematical formulae to argue for regular patterns in his data, quasi-deductive proofs that elevate these patterns to the status of a law, and the existing laws of combinatorics and probability to both inform the construction of his experiments and support the law-like regularity of his conclusions. Mendel's use of mathematical and quasi-mathematical formulae, operations, and laws added rigor to his biological arguments which appealed to his later supporters for whom, in 1900, mathematically describable laws were quickly becoming the gold standard for making arguments about evolution, heredity, and variation.

Interestingly, however, whereas Mendel's use of mathematics appealed to later audiences, this approach provoked an adverse reaction in his contemporaries who thought that Mendel was being presumptuous. In answer to the second question, I intend to make the case that, at least in part, Mendel's use of mathematical formulae and principles to make his arguments was rejected by his contemporaries because they believed that inherited characters were variable not stable and, therefore, could not be described using

²⁴ Combinatorics is described under the definition for "combinatorial" as, "relating to the arrangement of, operation on, and selection of discrete mathematical elements belonging to finite sets or making up geometric configurations." ("Combinatorial" *Merriam-Webster's Collegiate Dictionary*. 10th ed. 1998.)

a discrete mathematical system such as probability and combinatorics. In addition, they also objected to Mendel's use of mathematical formulae and principles to suggest that his results represented deductive laws of inheritance because his results only applied to seven trait pairs on a handful of pea species.

The tension created by Mendel's reliance on mathematics to express and argue his position, in addition to the non-conformity of his results to the traditionally accepted description of inheritance, provide important clues as to why for more than 30 years Mendel's work went unrecognized. Though Mendel makes a valiant rhetorical effort to persuade his audience to accept his conclusions and his methodology, the silence of 30 years bears witness to the failure of his persuasive approach with his intended readers regardless of the verity of his procedures and results.

The Hybridist Tradition

In order to understand the traditional sources of Mendel's argument and why Mendel's use of quantification, mathematical formulae, mathematical operations, and mathematical principles to make that argument represents a radical departure from his predecessors, it is important to know something about their arguments, beliefs and practices. Although experimentation in hybridization was considered as far back as 1694 by Rudolph Jacob Camerer²⁵ in his work *Über das Geschlecht der Pflanzen (About the Gender of Plants.)*, systematic research in hybridization, the crossing of plants considered to be different varieties and species, began in the mid-eighteenth century with the work of Joseph Gottlieb Kölreuter (1733-1806) and was carried on in the nineteenth century by

²⁵ Better known under his Latinized name Camerarius.

Carl Friedrich von Gärtner (1772-1850). In general, these eighteenth- and nineteenth-century researchers dealt with questions regarding the variation or fixity of natural forms and the physiological process by which either variety or homogeneity was transferred from one generation to the next. Specifically, they wanted to know how much parents contributed of their characters to their offspring, whether their contributions were qualitatively similar, how the characters of the parents were carried into the offspring, and how they developed in the offspring.

Joseph Gottlieb Kölreuter

Joseph Gottlieb Kölreuter was born on April 27, 1733 in the Sawbian village of Sulz in the Black Forest region of Southwest Germany. His early hybridization experiments were conducted in his hometown of Sulz, in the garden of Achatius Gärtner in the nearby town of Calw as well as in St. Petersburg, Berlin, and Leipzig. Later, after 1764, all of his work was done at Karlsruhe where he was employed as a professor of natural history (Roberts 35-36).

Kölreuter's four-volume work, which included *Vorläufige Nachricht von einigen das Geschlecht der Pflanzen betreffenden Versuchen und Beobachtungen* (Preliminary Report about some Observations and Experiments Regarding the Gender of Plants) and its three *Fortsetzungen* (*Continuations*), which were written in the above mentioned places from 1761 to 1766, became the foundational writings for the plant hybridists of the nineteenth century. In these works, Kölreuter sets down many of the central principles of

hybridization held also by Mendel, though his own work offers important deviations from them.

In this work, Kölreuter describes experiments in which he creates hybrids by crossing a total of 13 genera and 54 species over a six-year period, including his famous crossing of two different varieties of tobacco *Nicotina paniculata* and *Nicotina rustica* which represents the first recorded experiment in artificial plant hybridization (Roberts 36). These experiments were conducted by Kölreuter as a means of exploring reproduction in plants, especially the manner in which and the degree to which male and female reproductive material contributed to the formation of a new organism. Hybridization provided an ideal method for conducting such an investigation because the contrast between the features of the parents offered a means by which the experimenter could determine the nature of each of their contributions. If hybrid characters favored one or the other of the parents, Kölreuter could assume that the generative force of that parent was stronger than the other.

What he found, however, in the crosses of tobacco was that the hybrid always seemed to be a form intermediate between the two parents. The outcome of these crosses led Kölreuter to reconfirm Camerer's position that plants do have different sexual components which contribute material to the reproductive process. It also suggested to him that not only do both the male and female parents contribute to the production of offspring, but also that each of their contributions is approximately equal and blended as the hybrid seems to be a perfect intermediate of maternal and paternal characteristics (Roberts 42-43).

Despite the answers that his initial experiments seemed to offer in terms of the role of the parents in reproduction, they also led Kölreuter to an uneasy contemplation of the “naturalness” of the crosses he had created and whether or not they revealed a change in the essence of the two species involved. As a proponent of the philosophy of essentialism, inherent in the Christian doctrine of special creation, which held that each species was a unique and unalterable act of creation, Kölreuter was troubled by his ability to create a hybrid type challenging the notion that species were stable. When the first generation of crosses matured, however, to Kölreuter’s relief, they were found to be completely infertile. This confirmed to his satisfaction that, though they might be crossed, a divine barrier existed which would not let the transformation continue further and, therefore, disturb the essence of the species (45).

As he progressed in his experiments, however, Kölreuter was faced with more challenges because he discovered that crosses of plants which were more closely related than the tobacco species in his first experiment did produce viable seeds. After closer inspection of the number of seeds, though, Kölreuter concluded that these crosses were always less fertile than their parents, a fact which seemed, at least partially, to shore up the essentialist nature of the species. In addition, his convictions were reconfirmed by further experiments in which he back-crossed hybrids with their original parent species. These crosses resulted in plants most of which were similar to one of the two original species, a few of which were related to the form of the other parent species, and a few of which displayed faint mixed traces of both parents. The reversion of most plants to their parental forms in these experiments led Kölreuter to conclude that this, too, was evidence of a barrier in nature against the mixing of specific essences (Mayr 644).

From the results of these back-crosses, Kölreuter developed a theory of the process of inheritance in plants which held that each species or variety involved in hybridization expressed itself in different degrees, depending on the strength of a particular species' essence. (This concept of dominance appears in Mendel's work, though it is a characteristic of a particular genetic trait rather than a species' essence.) Those species with strong essences would be more present in back-crosses than those with weaker essences because their reproductive material asserted itself more vigorously or in greater amounts than the material provided by the weaker species.

Since these materials are of different sorts, or in their essence are different from each other, it is easy to comprehend that the strength of the one must be different from the strength of the other. From the union and commingling of these two materials, which occurs in the most intimate and orderly manner, according to a definite relationship, there originates another, which is of an intermediate sort.... All the movements and changes, which from the embryo to the time of flowering, take place in every masterpiece of nature, appear to be directed simply to the great work of reproduction. They all aim at gradually liberating that compound material upon which they are based, and at dividing it again into the two original ground materials; or, to speak more properly, to bring these latter themselves into a complete, and, especially from the one side, into masses of unlike size than were demonstrated from the preceding reproduction. (quoted in Roberts 47-48)

Here Kölreuter explains that reproduction is a process of combining and separating the essences of species. In the initial stage, the two essences are “commingled” and create an “intermediate material” for the hybrid plant. During the plant’s process of development, however, the commingled essences are separated and expressed in what would now be called the phenotype,²⁶ the outward physical appearance of the plant, according to the original strength of their essence. The result is a hybrid offspring (F1) which resembles both the parental species in keeping with the degree of vitality of their essences.

Kölreuter’s reasoning about the combination and expression of reproductive elements in the case of closely related species relies on an argument from antithesis. First, he believes that each species has its own essence which is unique from and contrasts with the essence of other species. Second, though these species combine in the process of hybridization, they separate out again in the phenotype where each exhibits dominion over one specific aspect of the phenotypic expression.

Kölreuter’s work provided the essential material and methods for all of the hybridists who came after. Most importantly they took up his ideas 1) that hybrids represented an equal or near equal commingling of parental species essences, 2) that hybrids of species are generally infertile with closely related forms showing a higher degree of fertility, 3) that in the offspring of fertile hybrids, a few of the offspring reverted to the parental hybrid form while most reverted to the grandparent forms in the original cross, and 4) that the degree to which traits are expressed in the physical

²⁶ Neither Kölreuter nor Gärtner had a theory which included the genotype, the collection of both expressed and hidden characters.

characteristic of hybrids is directly related to the strength of species' essences which are commingled.

In addition to his ideas, later researchers also adopted many of Kölreuter's methods for conducting experiments, drawing conclusions, and presenting results in writing. In the bulk of his work, Kölreuter worked with crosses between what were at the time labeled distinct species, e.g. *Nicotiana paniculata* X *Nicotiana rustica*, though the subsequent fertility of crosses led him to believe in some cases that certain species might be better described as varieties. From these carefully recorded crosses, he gathered both qualitative and quantitative data on the appearance of hybrids and their relations to their parents from which he inductively drew conclusions about the manner in which and degree to which traits were passed on from parents to offspring by their expression within the offspring.

The results of Kölreuter's crosses were presented graphically using natural language, brackets, and the symbols for male ♂ and female ♀.²⁷ He also employed quantitative measurements in order to catalogue the generations in crosses and to present comparative quantitative measurements of various physical features of the parents and their hybrid offspring.²⁸ The latter application of quantification represented an important component of Kölreuter's case because, by juxtaposing the measurements of parent plants and their offspring in a chart format, he was able to show that in some cases hybrid crosses were intermediate forms of their parents and in other cases one parent's character dominated. For the most part, however, Kölreuter's work did not rely heavily on

²⁷ See Appendix F

²⁸ See Appendix G

quantification nor make an effort to deduce any laws of regularity based on mathematical principles.

Carl Friedrich von Gärtner

In the German village of Calw, in the same garden where Kölreuter had conducted some of his hybridization experiments, one of the greatest supporters of his work, Carl Friedrich von Gärtner, undertook his own research program of crossing hundreds of species and varieties in the early nineteenth century (Olby 49). A physician by trade and son of a distinguished botanist, Joseph Gärtner (1732-91),²⁹ Carl Gärtner did 25 years worth of meticulous hybrid crossing experiments on his own³⁰ which he carefully recorded. The initial motivation for carrying out these experiments is not clear. In fact, Gärtner's results might have passed into obscurity altogether if not for a prize offered by the Dutch Academy of Sciences at Haarlem for essays answering the query, "What does experience teach regarding the production of new species and varieties, through the artificial fertilization of flowers of the one with pollen of the other, and what economic and ornamental plants can be produced and multiplied in this way?" (Roberts 167)

In October of 1835, Gärtner learned of the prize and sent the academy a brief sample of his work. Because he had not formally compiled all of his results, he asked for

²⁹ Author of an authoritative work on the seeds and fruits of plants *De Fructibus et Seminibus Plantarum* (1788)

³⁰ Based on the estimate given in Robert's for the number of years Gärtner spent on these experiments and the fact that the results were published in 1837, I assume that these experiments began around 1811 or 1812. (Roberts 168)

an extension to complete his work. The extension was granted and two years later he presented them with a two-hundred page memoir which in 1849 was published as a monograph under the title *Versuche und Beobachtungen über die Bastarderzeugung im Pflanzenreich* (Experimentation and Observation of the Creation of Hybrids in the Plant Kingdom). In this monograph Gärtner details nearly 10,000 separate experiments in crossing among 700 species belonging to 80 different genera of plants (168).

From these experiments Gärtner concludes, as Kölreuter had, that the essences of the different species are commingled in the hybrid; however, unlike Kölreuter, who believed in only one true hybrid form,³¹ Gärtner divides hybrids, based on the nature of their commingling, into three categories: 1) intermediate, 2) commingled, and 3) definite. Hybrids which expressed in equal amounts pure traits from each of their parent species were identified as members of the *intermediate* class of hybrids because “a complete balance occurred of both fertilizing materials, in respect to either mass or activities” (168). In this definition, Gärtner describes what he understands are two important factors which affect the degree of expression of a particular species’ essence in the hybridized offspring: 1) the *mass* or amount of a particular species essence present in an offspring and 2) the *activity* or potency of the essence present.

What Gärtner argues is that the presence of more or less mass or activity on the part of one species’ essence results in the expression of a lesser or greater number of that species’ characters in the phenotype of the hybrid offspring. Like Kölreuter, Gärtner believes in the expression of different aspects of the two species’ essences in different parts of the phenotype, but makes no attempt to experiment on specific traits because he sees traits as expressions of different facets of the unified fabric of a species’ essence

³¹ The form that Gärtner identifies as “intermediate.”

rather than as a mosaic of separate units. He describes the antithesis of essences and their expression in different domains of the phenotype when he writes,

In the formation of simple hybrids, as in sexual reproduction in general, two factors are active. This unlikeness of activity, flowing from the specific differences of species, expresses itself through the more pronounced or the weaker manifestation of the individual parental characters in the different parts of the hybrid. (Quoted in Roberts 171)

Whereas the intermediate form is created by an exact balance of contributions from one parent or the other, the *definite hybrid*, according to Gärtner, is one, “among which the resemblance of a hybrid to one of its parents, either to the father or the mother, is so marked and preponderating that the agreement with the one or with the other is unquestioned” (169). Here, Gärtner suggests that when the mass or activity of one of the species is greater than the other, the pure features of the predominant species are preponderant in the form of the offspring. Though Gärtner’s concept of preponderance appears similar to Mendel’s notion of dominance, it is different in the sense that Gärtner has no recessive counterpart. There is no sense of what happens to the reproductive material for a particular feature contributed from the non-dominant parent.

Gärtner’s sense that dominance is an all or nothing phenomenon in which features are inherited in pure form or eliminated entirely is clear in his definition of the second category of hybrid, the *commingled hybrid*, and in his understanding of the distribution of characters in second generation crosses. In the commingled hybrid, Gärtner explains,

Now this and now that part of the hybrid approaches more to the maternal or to the paternal form, whereby, however, the characters of the parents, in their transference to the new organism, never go over pure, but in which the parental characters always suffer a certain modification. (Roberts 168)

Though there is no statement here about the forces which might affect modification in the parental characters, there is no hint that this modification of character is influenced by the mixing of characters from one parent with the other. Instead, there is a sense that the characters are modified before they are blended in pollination.

In his discussion of back-crosses, however, Gärtner seems to suggest that the modification in the pure characters might have something to do with the presence of the essence of the different parent species. However, again, the modification is not one in which different contributions from each of the parents are mixed to create a new character which is not purely one or the other. Instead, the presence of each of the species' essences has a modifying influence on the manner in which the dominating features of the other is expressed.

[The fundamental ground material of the hybrid] behaves differently in the second and in the further stages of breeding, where, on account of the different nature of the two factors of the hybrids in the succeeding fertilizations, an altered, shifting, variable direction in type formation enters into the arising varieties (169).

In their work, Gärtner and Kölreuter raised important questions about the processes of reproduction in plants. They asked: What happens when two species are combined? And they answered: In the process of hybridization, the essences of two species combined and separated out with each species expressing itself in a particular domain of the offspring. They queried: Why do the characters of one species predominate in some cases and not in others? And they responded: Because the mass or activity of one species in the case of a particular character was greater than the other. Gärtner wondered: Why are features not always purely expressed in offspring and in later generations? And he theorized: That either outside influences or influences from the presence of the essence of the other species affected the pure expression of characters.

Mendel's understanding of the fundamental precepts of plant hybridization theory and his primary argument for the stability of traits came from the works of Kölreuter and Gärtner. However, the impetus for his research was the result of questions which still lingered in the minds of hybridists regarding the quality and quantity of the contribution that each different parental species made to the hybrid. Though Gärtner and Kölreuter had theorized about why hybrid forms had a mixture of characters from either parent which, in certain cases, varied in different degrees, researchers still sought a clearer understanding of what accounted for such differences in the expression of characters.

The Origin of Mendel's Experiments

In 1856 Mendel began eight years worth of hybridization experiments with peas. Although the results of these experiments are recorded in "Experiments in Plant

Hybridization,” there is no discussion within the text, or in any other document for that matter, which explains why Mendel set out to do these experiments in the first place.

There are two theories, however, which have been widely discussed by Mendel scholars (Henig 2000, Iltis 1932, Orel 1984).

First, there is the belief that the experiments were inspired by Mendel’s second failure in 1856 of the examination for high school teachers. The legend goes that in the examination Mendel had disagreed with one of his examiners, Eduard Fenzl. Fenzl was the director of the Vienna botanical gardens and was firmly of the opinion that, “The plant embryo resided, microscopic but entirely preformed, in the pollen and passed to the ovary through the pollen tube. All it needed to do was grow; the female part of the plant offered nothing more than an environment that made growth possible” (Henig 62).

Conversely, Mendel believed that there was no a priori existence of a complete embryo in the male. Instead, he held that the embryo was created only with the coming together of the male and female reproductive material. As the story goes, this position became a point of contention in the examination. Mendel chose to fail rather than capitulate to Fenzl’s point of view.

Angered by the debate and his second failure, Mendel presumably returned to Brunn and set to work to prove, once and for all, that his position was correct. The series of carefully planned breeding experiments with peas, therefore, was his defiant scientific gesture to Fenzl and the spermist theory of reproduction.

Although the importance of Mendel’s run-in with Fenzl as a source of inspiration is entertained by some Mendel scholars (Henig 2000), ignored by others (Olby 1966 and Orel 1984), and dismissed altogether by still others (Iltis 1923), that a theory of

fertilization is at the heart to his work cannot easily be ignored. Though there is no telling whether Mendel's dispute with Fenzl, even if it did happen, inspired him to develop his hybridization experiments, I will argue that Mendel certainly was aware that understanding this process was key to supporting his theory of the distribution of characters.

In fact, the work of Max Ernst Wichura (1817-66) on the hybridization of willows, cited by Mendel in his "Experiments in Plant Hybridization," is the first text in the hybridization literature which argues that during amphimixis, the joining of the male and female reproductive material, each of the parents must contribute an equal amount because in reciprocal crosses the traits of each of the parents seemed to segregate out in a regular pattern (Roberts 181-2).

In the analysis of Mendel's text, I will show that, like Wichura, Mendel also argues that equality must exist in the combination of reproductive material because of the mathematical regularity in the offspring of hybrids. In fact, this position, I will argue, is crucial in designing his hybridization experiments and, therefore, must have been clear to Mendel and an important influence on him either before or in the early stages of his breeding experiments.

A second theory about why Mendel undertook his program of hybridization, one that has the broadest support among historians, is that he was interested in providing support for the concept that populations of organisms varied over time. This theory is suggested by Iltis who writes,

Although since the famous dispute between Cuvier and Geoffroy Saint-Hilaire the doctrine of the permanence of organic forms seemed to have definitely gained the upper hand over any theory of their progressive evolution, all far-seeing biologists were still cudgeling their brains over the problem, “How is the extraordinary multiplicity of living forms to be explained?” Mendel’s interest in his ornamental flowers may first have been aroused by... pure delight... but the peas and the other plants... were in the end to furnish him with answers to serious problems. (106-07)

The theory is also supported by Henig:

Maybe Mendel set out, in his pea experiments, to confirm his idea of “perennial” progressive change as a driving force for the appearance of new species from the old. If he designed his experiments correctly, he could lend empirical support to the theories of two botanists he had come most to admire: Nägeli and Unger. (64)

During the 1840s and 1850s a radical biological movement developed in the work of Franz Unger, Karl Nägeli, and others which rejected Gärtner and Kölreuter’s position that species were fixed categories and embraced Jean Baptist Lamarck’s (1744-1829) position in *Philosophie Zoologique* (1809) that species could change over time. Franz Unger (1800-70) was Mendel’s professor of physiology and plant paleontology in Vienna from the fall of 1852 to the spring of 1853. In his weekly “Botanical Letters” column in the *Vienna Times* (*Wiener Zeitung*) in the winter of 1856 Unger wrote, “Who can deny

that new combinations arise out of this permutation of vegetation, always reducible to certain law combinations, which emancipates themselves from the proceeding characteristics of the species and appear as a new species” (quoted in Henig 63).

These words of Mendel’s instructor are companions to the sentiments of Carl von Nägeli who, in a work published in 1856 *Individuality in Nature with especial Reference to the Vegetable Kingdom (Die Individualitat in der Natur mit besonderer Berucksichtigung des Pflanzenreiches)*, writes,

Like natural phenomena in general, species cannot persist in complete repose. Just as the offspring of the first individual were a little different from that individual, so also must the germs which engendered them diverge to some extent from those out of which they themselves originated. A process of change must be perennially at work, and this change cannot fail, in the end, to bring about the disappearance of the species or its transition into another. (Quoted in Iltis 186)

Although Nägeli had no direct connection to Mendel until the year after Mendel’s hybridization experiments were in print, Unger was a proponent of Nägeli’s work and most likely referred to him in his lectures (Orel 40).

In addition to there being evidence that there were academic influences in Mendel’s life that might have encouraged him to undertake experiments to prove variation in organisms, Mendel himself suggests in the opening of “Experiments in Plant Hybridization” that, at least in part, his experiments had been undertaken to contribute to the discussion regarding the variability of the species.

It requires indeed some courage to undertake a labor of such far-reaching extent; this appears, however, to be the only right way by which we can finally reach the solution of a question the importance of which cannot be overestimated in connection with the history of the evolution of organic forms.^[32](Mendel 2)

Though there seems to be sufficient evidence to support the view that Mendel undertook his experiments to support Unger and Nägeli's position on the variability of species, a crucial piece of evidence has been overlooked in these assessments. In the conclusion of the text, Mendel argues that species are in most cases *not* variable.

He [Gärtner] perceives in the complete transformation of one species into another indubitable proof that species are fixed within limits beyond which they cannot change. Although this opinion cannot be unconditionally accepted, we find on the other hand in Gärtner's experiments a noteworthy confirmation of that supposition regarding variability of cultivated plants which has already been expressed.^[33]

³² Es gehört allerdings einiger Muth dazu, sich einer so weit reichenden Arbeit zu unterziehen; indessen scheint es der einzig richtige Weg zu sein, auf dem endlich die Lösung einer Frage erricht werden kann, welche für die Entwicklungs-Geschichte der organischen Formen von nicht zu unterschätzender Bedeutung ist. The original German is from a reprint of Mendel's original text: Gregor Mendel, *Versuche über Pflanzen-Hybriden*, ed. J. Cramer and H.K. Swann (1865; New York: Hafner Publishing, 1960) 4. All subsequent footnotes containing the original German will be from this text.

³³ On page 32 of "Experiments" Mendel writes, "But nothing justifies the assumption that the tendency to the formation of varieties [in cultivated plants] is so extraordinarily increased that the species speedily lose all stability, and their offspring diverge into an endless series of extremely variable forms." (Allein nichts berechtigt uns zu der Annahme, das die Neigung zur Varietätenbildung so ausserordentlich gesteigert werde, dass die Arten bald alle Selbstständigkeit verlieren un ihre Nachkommen in einer endlosen Reihe höchst veränderlicher Formen auseinander gehen. (*Versuche*, 36))

Among the experimental species there were cultivated plants... and hybrids between these species lost none of their stability after four or five generations. [³⁴] (41)

Though Mendel's results failed in the end to support Nägeli and Unger's belief in variability, this failure does not itself preclude the possibility that Mendel began his experiments with the intention of supporting their position. In the "Introductory Remarks" section of his text, however, Mendel suggests otherwise by citing as his precursors Gärtner, Kölreuter, and Wichura, all of whom believed in the fixity of the species. The work and ideas of Unger, Nägeli, and Darwin³⁵ are not mentioned here or anywhere else in the text.

The striking regularity with which the same hybrid forms always reappeared whenever fertilization took place between the same species induced further experiments to be undertaken, the object of which was to follow up the developments of the hybrids in their progeny.

³⁴ Es [should read "Er"] sieht in der vollendeten Umwandlung einer Art in die andere den unzweideutigen Beweis, dass der Species feste Grenzen gesteckt sind, über welche hinaus sie sich nicht zu ändern vermag. Wenn auch dieser Ansicht eine bedingungslose Geltung nicht zuerkannt werden kann, so findet sich doch anderseits in den von Gärtner angestellten Versuchen eine beachtenswerthe Bestätigung der früher über die Veränderlichkeit der Culturpflanzen ausgesprochenen Vermuthung.

Unter den Versuchsarten kommen cultivirte Gewächse vor... und auch diese hatten nach einer 4 bis 5 maligen hybriden Verbindung nichts von ihrer Selbstständigkeit verloren. (*Versuche* 46-47)

³⁵ Of Darwin Mendel has been reported as saying that he was, "greatly interested in the ideas of evolution, and was far from being an adversary of the Darwinian theory." However, he is said to have stated that, "there was still something lacking" (Orel 71).

To this object numerous careful observers, such as Kölreuter, Gärtner, Herbert, Locoq, Wichura, and others, have devoted a part of their lives with inexhaustible perseverance.³⁶ (1)

In this section of the introduction, Mendel puts his work in the context of those hybridists who have been struck by the “regularity” and “reappearance” of traits during hybridization, not by their variability or disappearance. Though he is reverent regarding the conclusions of these earlier researchers, Mendel expresses dissatisfaction with the methods by which they came to their conclusions. He argues that in order to rigorously confirm or refute the stability of species, a more precise quantitative program of breeding experiments has to be carried out to ascertain what patterns, if any, might exist in the distribution of traits.

Those who survey the work done in this department will arrive at the conviction that among all the numerous experiments made, not one has been carried out to such an extent and in such a way to make it possible to determine the number of different forms under which the offspring of the hybrids appear, or to arrange these forms with certainty according to their separate generations, or definitely to ascertain their statistical relations.³⁷ (1-2)

³⁶ Die auffallende Regelmässigkeit, mit welcher dieselben Hybridformen immer wiederkehrten, so oft die Befruchtung zwischen gleichen Arten geschah, gab die Anregung zu weiteren Experimenten, deren Aufgabe es war, die Entwicklung der Hybriden in ihren Nachkommen zu verfolgen.

Dieser Aufgabe haben sorgfältige Beobachter, wie Kölreuter, Gärtner, Herbert, Lecocq, Wichura u.a. einen Theil ihres Lebens mit unermüdlicher Ausdauer geopfert. (*Versuche* 3)

³⁷ Wer die Arbeiten auf diesem Gebiete überblickt, wird zu der Ueberzeugung gelangen, dass unter den zahlreichen Versuchen keener in dem Umfange und in der Weise durchgeführt ist, dass es möglich wäre, die Anzahl der verschiedenen Formen zu bestimmen, unter welchen die Nachkommen der Hybriden

Unlike his predecessors Mendel was uniquely situated to develop the quantitative hypothetico-deductive experimental design that was required to ascertain the number and arrangement of forms over many generations of breeding. The reason for his success is in part the result of Mendel being a hybrid himself: part biological scientist, part physical scientist, and part mathematician.

Mendel the Mathematician and Physical Scientist

A brief foray into Mendel's personal history reveals the extent of his interest and training in the physical sciences and mathematics. As a student at the University of Vienna from 1851-53, Mendel enrolled in courses with the great physicist Christian Doppler and with the renowned mathematician Andreas von Ettingshausen. In fact, Mendel's interest in mathematics and the physical sciences was so pronounced that it made up, according to Alain Corcos and Floyd Monhagan (in their book *Gregor Mendel's Experiments on Plant Hybrids*) at least half of his academic schedule (24).

Along with learning the subject matter of mathematics and physics, Mendel also had an education in how to apply the knowledge in these fields to investigations of the natural world. He was chosen by Doppler as an assistant demonstrator at the Physical Institute at the university. In this capacity, he learned how to perform experimental demonstrations of various physical phenomena. This practical education in the methodology of experiments in the physical sciences was employed by Mendel in the

aufzutreten, dass man diese Formen mit Sicherheit in den einzelnen Generationen ordnen und die gegenseitigen numerischen Verhältnisse feststellen könnte. (*Versuche* 3-4)

careful planning and design of his *Pisum* experiments, which, unlike traditional hybrid experimentation that relies on a process of first compiling and cataloguing information and then drawing conclusions based on these observations, employs the Newtonian hypothetico-deductive approach where a hypothesis is given and experiments are then carefully designed to prove the truth or falsity of that hypothesis.

In addition to the methodology of experimentation, Mendel was also introduced to the use and importance of mathematics as a tool for predicting and modeling physical phenomena. At the time of his studies, the mathematics of statistics and probability had become an important part of this area of knowledge (Orel 31). Orel suggests that Mendel might have learned something of the subject from Doppler, who had published a mathematical textbook³⁸ in 1844 with a chapter on, “combinatorial theory and basic principles of probability calculation” (31).

Orel also suggests that Mendel could have gained some understanding of the theoretical application of probability to natural philosophy from the work of Joseph Johann von Littrow (1781-1840) with whose small volume, *Probability Calculation as Used in Scientific Life* (1833), Mendel was also familiar. In the text, Littrow discusses the application of Laplace’s theory of probability to all phenomena arguing that, “the relationships of all phenomena in nature seem at first completely random; but the greater the number of those phenomena that are considered, the closer they approach to certain constant relationships” (32). Orel explains that Mendel relied on the book as a resource for compiling meteorological records and for making meteorological forecasts of the

³⁸ Orel does not provide the name of the textbook. I assume, however, that he is referring to *Arithmetik und Algebra: mit besonderer Rücksicht auf die Bedürfnisse des practischen Lebens und der technischen Wissenschaften: nebst einem Anhang von 450 Aufgaben* published in 1844.

weather in Moravia and that he may have also applied the methodology developed from this experience to his work with peas (32).

Although mathematical symbolization for the purpose of quantifying and then comparing experimental results were used by Kölreuter in his hybrid experiments, there was neither an effort on his part to use mathematical principles as a source of invention for his arguments, nor to make deductive arguments from those principles. Mendel's background in probability and his work with meteorology, however, suggests that even before he set out on his experiments with pea hybridization he may have been relying on established mathematical principles as aids for understanding natural phenomena.

The Role of Quantification, Combinatorics, and the Principles of Probability in Making Mendel's Case

With no primary record by Mendel of how he developed his arguments in "Experiments in Plant Hybridization," the only way to understand the role that quantification, combinatorics and the principles of probability played in developing his case for 1) the regularity of the distribution of characters over generations, 2) the particulate non-interacting nature of characters, and 3) the equality of the contributions of parental characters in offspring is by closely examining the arguments themselves in the text of the "Experiments." What this analysis reveals is that though Mendel's arguments have their source partially in the positions and methods of previous hybridists, they also rely for support on quantified data and on the principles of combinatorics and probability

to elevate the status of his findings from a series of observations to a set of deductive natural laws.

In the first five sections of Mendel's text following the "Introductory Remarks" ("Einleitende Bemerkungen") Mendel does exactly what he claims he will do in his introduction: 1) "determine the number of forms under which the offspring of hybrids appear" and 2) "arrange these forms with certainty according to their separate generations." His work in this section is very much informed by the traditional practices of hybridization; however, Mendel adds precision to his hybrid experiments by purposely choosing specific features to observe, carefully ensuring that plants are not accidentally pollinated, and precisely describing the relationship between different forms in different generations quantitatively.

In the first two sections following the introduction, Mendel provides his readers with a careful accounting of and rationale for both the materials and the methods of his experiments. He begins the first section, "Selection of Experimental Plants," by describing the three basic criteria he employed when choosing and controlling his experimental subjects.

1. [That they] Possess constant differing characters.
2. That the hybrids of such plants must, during the flowering period, be protected from the influence of all foreign pollen, or be easily capable of such protection.
3. The hybrids and their offspring should suffer no marked disturbance in their fertility in the successive generations.³⁹ (2)

³⁹ Die Versuchspflanzen müssen nothwendig

Of these criteria, the one which is unique to Mendel's work is the first one. It is unlike the criteria of earlier hybridists because 1) it begins experimentation from the position that a species is made up of a mosaic of separate features, and 2) it seeks to ensure that those features are pure or constant. These concepts are extremely important to Mendel's case. The idea that traits are separate elements is important because it allows Mendel to clearly identify a characteristic of a plant as a single unit and follow that unit from one generation to the next. The traditional beliefs held by Kölreuter and Gärtner, that all of the characters, or describable features of a plant, were part of a species essence and, therefore, had to be considered en toto, made describing the assortment of characters extremely difficult because it was impossible to follow all of the features at once from one hybrid cross to the next.

In addition to positing separate characters, the first criterion also requires that the features to be experimentally bred remain pure. This criterion is crucial to Mendel's arguments because if he is to prove that traits remain stable over many generations he must begin his experiments with traits that do not vary before hybridization. In addition, it provides evidence that Mendel was designing his experiments around a preconceived position that heredity could be described by mathematical principles, which require regularity. In order to ensure the stability of the traits in the seeds he was breeding, Mendel explains he has allowed the plants to self-breed. Then, after examining the offspring, he eliminated any plants which did not breed true (3).

-
1. Constant differierende Merkmale besitzen
 2. Die Hybriden derselben müssen während der Blütezeit vor der Einwirkung jedes fremdartigen Pollens geschützt sein oder leicht geschützt werden können.
 3. Dürfen die Hybriden und ihre Nachkommen in den aufeinander folgenden Generationen keine merkliche Störung in der Fruchtbarkeit erleiden. (*Versuche 5*)

Though the first criterion is perhaps the most innovative of Mendel's experimental parameters, the second and third criteria are no less important. In the case of the second parameter, Mendel reveals the importance of protecting the experimental plants from foreign pollen. He explains that if the experimental subjects are open for fertilization the results of the experiment will not be accurate. In order to ensure that uncontrolled fertilization does not occur, Mendel takes two precautions. First, Mendel uses peas, *Pisum*, because they self-fertilize and second, he follows the time honored practices of emasculating and crossbreeding plants just before self-fertilization can occur.

The final criteria, that species whose hybrids are fertile must be crossed, is important because without an observable series of all the possible permutations of offspring in successive generations it would be impossible to ascertain with any certainty whether or not there was any regularity in the expression of traits from one generation to the next. In *The Growth of Biological Thought*, Mayr explains that this particular parameter was also extremely fortunate for Mendel because, "Differences among intrapopulation variants are usually single-gene differences and display uncomplicated Mendelian segregation, while differences between species are often highly polygenetic and fail to segregate cleanly" (Mayr 713). Had Mendel chosen less fertile forms, which would have varied genetically to a greater extent, the regularity of the relationship in the different generations of hybrids would not have been so clearly observed. Mendel explains at the end of this section that after some preliminary crossings he was led to adopt 22 different varieties of the species *Pisum* as candidates for his crossing experiments (Mendel 4).

Once Mendel has introduced and explained his experimental criteria, he provides his readers with a list of the character traits that he intends to observe in his experiments.

There are seven characters including 1) the difference in the form of the ripe seeds (smooth/wrinkled), 2) the difference in the color of the seed albumin (pale yellow, bright yellow, or orange/intense green), 3) the difference in the color of the seed coat (white/grey, grey-brown, leather brown, or violet), 4) the difference in the form of the ripe pods (smooth/wrinkled), 5) the difference in the color of the unripe pods (light to dark green/vividly yellow), 6) the difference in the position of flowers (axial along the main stem/terminal at the end of the stem), and 7) the difference in the length of the stem (long 6-7ft./short $\frac{3}{4}$ -1 $\frac{1}{2}$ ft.) (5-6). Mendel then explains that these traits were chosen due to their high degree of visibility and the regularity of their behavior.

In some ways, this a priori choice of traits is a radical departure from previous hybridization experiments; however, in others it follows classical lines of thought regarding hybridization. It is radical on the one hand because, as mentioned before, most traditional hybridists did not think of the character traits as separate elements. Having separate antithetical traits, however, is necessary for Mendel if he is to complete the primary mission of his experiments which is to present a quantitative description of the distribution of traits within the organism over generations. This task is immensely simplified if 1) he can be assured that the traits being counted have been shown to breed true and 2) he can discriminate between one trait and another. In this way, Mendel's desire for quantified results influences the specific features of the experimental design making it amenable to the application of the formulae, operations, and principles of mathematics.

In other ways, of course, the choice of traits is very traditional. This is especially true of his interpretation of characters in terms of antitheses. This either/or distinction

reflects Kölreuter and Gärtner's impression that each species had a distinct essence and that that essence had a certain degree of dominance in the final form. Mendel does not question whether or not a species or variety might have three or four forms for the same feature. Instead, he is perfectly content to follow the traditional opinion of one species or variety, one form.

This mixture of the novel and the traditional in Mendel's experimental design seems to continue in the next three sections in which Mendel describes the results of crossbreeding experiments of the different varietal forms and the first and second generations created by allowing the hybrids to self-breed. In the first section, "The Forms of the Hybrids" [F1], Mendel describes the results of his crossing of varieties with the seven pairs of antithetical characters listed in the previous section. He begins by describing the conclusions made in previous work about the outcomes of crossing different species.

Experiments which in previous years were made with ornamental plants have already afforded evidence that the hybrids, as a rule, are not exactly intermediate between the parental species. With some of the more striking characters... the intermediate, indeed, is always to be seen; in other cases, however, one of the two parental characters is so preponderant that it is difficult, or quite impossible, to detect the other hybrid.⁴⁰ (Mendel 7)

⁴⁰ Schon die Versuche, welche in Früheren Jahren an Zierpflanzen vorgenommen wurden, lieferten de Beweis, dass die Hybriden in der Regel nicht die genaue Mittelform zwischen den Stammarten darstellen. Bei einzelnen mehr in die Augen springenden Merkmalen... wird in der That die Mittelbildung fast immer ersichtlich; in anderen Fällen hingegen besitzt das eine der beiden Stamm-Merkmale ein so grosses Uebergewicht, dass es schwierig oder ganz unmöglich ist, das andere an der Hybride aufzufinden. (*Versuche* 10)

In this explanation, Mendel acknowledges that in earlier crossbreeding experiments there have been two distinct results, either 1) the characters of the resulting hybrid offspring are intermediate between the characters of the parents, or 2) the character of one of the parents predominates. In the case of all of his experimental characters, the latter outcome has been the result (8).

Though it may seem coincidental that in all of the pairs of traits that Mendel chose for experimentation one form tends to predominate over the other, he makes it clear that this coincidence is no accident. The clear presence of one predominating form is necessary for quantifying the distribution of traits in different generations of crossbreeds. With a clear predominance of one or the other feature, it is possible for Mendel to categorize a form as form *x* or form *y* and tabulate the ratio of forms in each generation. If the forms were blended, this task would be impossible because there would be no reliable way of sorting offspring into one category or another, and, therefore, there would be no way of making quantitative comparisons. In a very real way, then, Mendel's desire for a mathematically based argument dictates the type of characters—those that do not blend—he uses as evidence for his theory.

For the character in a pair which predominates, Mendel uses the term *dominant* (dominirende) and for that character which disappears in the first hybrid generation he uses the term *recessive* (recessive).⁴¹ These names designate a different aspect of the relationship between the paired traits. This relationship remains the central focus in the rest of the paper. The actual qualitative differences—wrinkled and smooth etc.—lose

⁴¹ In der weiteren Besprechung warden jene Merkmale, welche ganz oder fast unverändert in die Hybride-Verbindung übergehen, somit selbst die Hybriden-Merkmale repräsentiren, als dominirende, und jene, welche in der Verbindung latent warden, als recessive bezeichnet. (*Versuche* 11)

purchase in all but the parts of the discussion in which he uses them to distinguish between the different character pairs present in bi- and tri-hybrid crosses.

Following this distinction between dominant and recessive, Mendel adds that it does not matter whether the character which is dominant comes from a male or female plant. This, he argues, is in keeping with the observations of Gärtner, who also believed that the dominance of a character was not linked to gender. This fact is important for Mendel to establish because it eliminates sexuality as a source of dominance.

After setting up all the parameters of the subjects used in the hybrid crosses, Mendel describes the result of two generations of self-crosses from the hybrids. Because of the way he has set up the experiments, it is possible for him to present the results quantitatively. In “The Generation from the Hybrids” [F2] (Die erste Generation der Hybriden), Mendel begins by explaining that the dominant traits from the previous generation appear, but the recessive traits, which had disappeared in the first cross, reappear as well and the dominant and recessive plants appear in an average proportion of three to one. After describing this trend, he provides his readers with the actual quantified results of his crosses. For example, Mendel describes the data obtained from observing the forms of the seeds generated from self-crossing *F1* hybrids which were the result of crossing parents *F0* with the dominant trait, round seed form, and the recessive trait, angular wrinkled seed form.

Expt. 1. Form of seed. – From 253 hybrids 7,324 seeds were obtained in the second trial year. Among them were 5,474 round or roundish ones and 1,850 angular wrinkled ones. Therefrom the ratio 2.96 to 1 is deduced.⁴² (Mendel 9)

Following quantitative descriptions of the first two experiments, Mendel stops to explain the manner in which the different dominant and recessive forms were distributed within individual plants. This aside is extremely important because with it Mendel makes the case from the probabilistic principle of large numbers that the ratio he is describing can only be discovered by looking at a large sample of plants. According to the principle of large numbers, the more observations/measurements investigators make of a particular phenomenon the closer they can get to the true value of that phenomenon. This principle is particularly important in the investigation of peas because the ratio is not exhibited in each example, i.e. if you split open a single pod there will not be three dominant round seeds and one recessive angular one. Instead, the observer needs to collect and quantify the instances of dominant and recessive seeds from many plants and add them up to arrive at the ratio (10).

Having explained the rationale driving his experimentation, Mendel proceeds to accumulate more empirical evidence to support his 3:1 ratio of dominant to recessive in the first generation of hybrid crosses. He reports the results from the crossing of all seven characters all of which, in varying degrees, approximate three to one.

After having made the case for his 3:1 ratio of dominant to recessive in the *F*₂ generation, Mendel prepares his reader for the argument he plans to make in the next

⁴² 1. Versuch. Gestalt der Samen. Von 253 Hybriden wurden im zweiten Versuchsjahre 7324 Samen erhalten. Darunter waren rund oder rundlich 5474, und kantig runzlig 1850 samen. Daraus ergibt sich das Verhältniss 2,96: 1 [the comma is used instead of a decimal point]. (*Versuche* 12)

section. He explains that of the three dominant forms in the ratio two of them are heterozygous,⁴³ i.e. they contain within their genetic pool one dominant and one recessive character (11-12). This distinction between pure or “parental forms” and mixed “hybrid forms” creates a second set of antithetical relationships, this time based on the genotype rather than the phenotype.

Just as the recessive trait could only be uncovered after the self-breeding of the hybrid generation *F1*, the existence of heterozygous dominant could only be revealed when the dominants of the *F2* generation were bred to reveal that some, the heterozygous dominants, had dominant and recessive offspring while others, the homozygous dominants, bred true. Using the collective number of instances of dominant and recessive offspring produced by self-breeding *F2*, Mendel arrives at a definite ratio of the number of hetero- to homozygous plants in *F2* and supports his previous 3:1 ratio of dominant to recessive (12).

As in the previous section, Mendel follows this conclusion with the quantified results of his experiments. Using these results, he accumulates evidence supporting his position. Then, he provides his reader with a summative quantitative expression of his findings, “The ratio of 3 to 1... resolves itself therefore in all experiments into the ratio 2:1:1 [two, in contemporary terms, heterozygous dominants, one homozygous dominant, and one homozygous recessive]” (13).⁴⁴ Following the accumulation and quantitative

⁴³ In order to differentiate heterozygous from homozygous plants in the “Experiments” Mendel uses the quasi mathematical letter combination *Aa* and the phrases “hybrid character” (hybrides Merkmale) and “hybrid forms” (Hybridformen). He contrasts these phrases with “dominant character” (dominirenden Merkmale) and “parental character” (Stamm-Character) which describe the homozygous forms. He never uses the term “heterozygote,” which was coined in 1902 by William Bateson the leading advocate of Mendel’s work in England. (“Heterozygote,” *Oxford English Dictionary* (Oxford: Oxford University Press, 2006) 27 Mar. 2006 <<http://dictionary.oed.com/>>.)

⁴⁴ Das Verhältniss 3:1... löst sich demnach für alle Versuche in de Verhältniss 2:1:1 auf. (*Versuche* 16)

description of his evidence, Mendel ends the section with a qualitative statement of what he is convinced is a regular pattern or law of heredity. The italicized presentation of the statement highlights its importance as a significant conclusion.

*It is now clear that the hybrids form seeds having one or other of the two differentiating characters, and of these one-half develop again the hybrid form, while the other half yield plants which remain constant and receive the dominant or the recessive characters [respectively] in equal numbers.*⁴⁵ (Mendel 13)

In the five sections following the introductory remarks, Mendel relies on a mixture of traditional experimental techniques and non-conventional mathematical practices and principles to develop his experiments and present his arguments. His experimentation with two antithetical forms and his methods for cross-pollination and emasculation of plants followed the tried and true practices of the hybridists before him. Also, his assumption that dominance was not an exclusive characteristic of either the male or the female plant was supported by Gärtner and other hybridists of the time.

However, the primary mission of Mendel's experiment, to provide a quantified description of forms, had not been embraced by other hybridists. This goal of quantification affected Mendel's experimental design such that it differed radically from his predecessors. First, the move towards quantification encouraged him to simplify the subject of observation from all of the characteristics of each contributing species/variety in the cross to particular traits of each. Second, it encouraged him to look only at

⁴⁵ Wird es nun ersichtlich, dass die Hybriden je zweier differirender Merkmale Samen bilden, von denen die eine Hälfte wieder die Hybridform entwickelt, während die andere Pflanzen gibt, welche constant bleiben, und zu gleichen Theilen den dominirenden und recessiven Character erhalten. (*Versuche* 17)

characters which, when crossed, resulted in offspring that clearly exhibited one trait or another. Without these simplifications, quantification would have been extraordinarily difficult, if not impossible.

In addition to quantification, there is also evidence in Mendel's methodological explanations that the principles of statistics and probability were driving his experimental design. The influence of probability is evident in his discussion of the appropriate quantity and quality of data for calculating ratios. In this explanation, he makes the case for his data collection practices based on the theory of large numbers, which suggests that the more data collected the closer the researcher can get to the true value of a particular phenomenon. In addition, it reveals Mendel's tendency to think of character ratios in terms of the population because he looks at the cumulative ratio of all traits in all plants, rather than of the specific ratio of traits in and between individual plants.

In the first five sections following the introduction, Mendel takes the initial steps in making his case for the non-variability of characters using mathematics. First, he quantifies the phenomenon, translating it into the correct form to be assessed mathematically. Second, he suggests a series of fixed quantifiable relationships between characters. These initial steps of moving from the qualitative to the quantitative domain do not require Mendel to rely heavily on mathematical notation or principles.

From Word into Law

Once his evidence is quantified and the relationship between the characters has been described both qualitatively and quantitatively, Mendel turns to the special topics of

combinatorics and probability to provide deductive arguments to support the claims made from inductive reasoning in the first part of the text: that the characters he is examining remain fixed over time and in a constant, quantifiable relationship to one another.

Further, he turns to the deductive principles of mathematics to support his contention that this relationship should be considered a *law* of nature. As a result of this movement from quantification to rational certification, the role of mathematics in making his arguments becomes much more visible and extensive in the text.

Mendel moves to the next phase of his argument in the section titled “The Subsequent Generations from the Hybrids” (“Die weiteren Generationen der Hybriden”). In the opening lines of this section, he makes it clear to his readers that his aim is to establish a law of heredity based on the results of his experiments. He writes, “The proportions in which descendents of the hybrids develop and split up in the first and second generations presumably hold for all subsequent progeny”⁴⁶ (Mendel 13).

The first step Mendel takes in making his case, that the relationships he has witnessed in his experiments represent a law of nature, is to transform the dominant and recessive character traits and their quantitative relationship to one another into a single quasi-mathematical expression: $A+2Aa+a$.

If A be taken as denoting one of the two constant characters, for instance the dominant, a , the recessive, and Aa the hybrid form in which both are conjoined, the expression

$$A+2Aa+a$$

⁴⁶ Die Verhältnisse, nach welchen sich die Abkömmlinge der Hybriden in der ersten und zweiten Generation entwickeln und theilen, gelten wahrscheinlich für alle weiteren Geschlechter. (*Versuche* 17)

shows the terms in the series for the progeny of the hybrids of two differentiating characters⁴⁷. (14)

After the initial presentation of this mathematical expression describing the ratio of traits in a hybrid generation, Mendel, realizing perhaps that his audience might object to his mathematical transformation, attempts to win their consent by showing that his expression provides deductive proof for an observation made by Kölreuter that had long been an accepted commonplace for hybridists.

The observation made by Gärtner, Kölreuter, and others, that hybrids are inclined to revert to the parental forms, is also confirmed by the experiments described.... If an average equality of fertility in all plants in all generations be assumed, and if, furthermore, each hybrid forms a seed of which one-half yields hybrids again, while the other half is constant to both characters in equal proportions, the ratio of numbers for the offspring in each generation is seen in the following summary.⁴⁸ (Mendel 14)

⁴⁷ Bezeichnet A das eine der beiden constanten Merkmale, z. B. das dominirende, a das recessive, und Aa die Hybridform, in welcher beide vereingt sind, so ergibt der Ausdruck:

$$A + 2Aa + a$$

die Entwicklungsreihe für die Nachkommen der Hybriden je zweier differirender Merkmale. (*Versuche* 17)

⁴⁸ Die von Gärtner, Kölreuter und Anderen gemachte Wahrnehmung, dass Hybriden die Neigung besitzen zu den Stammarten zurückzukehren, ist auch durch die besprochenen Versuche bestätigt.... Nimmt man durchschnittlich für alle Pflanzen in allen Generationen eine gleich grosse Fruchtbarkeit an, erwägt man ferner, dass jede Hyride Samen bildet, aus denen zur Hälfte wieder Hybriden hervorgehen, während die andere Hälfte mit beiden Merkmalen zu gleichen Theilen constant wird, so ergeben sich die Zahlenverhältnisse für die Nachkommen in jeder Generation aus folgender Zusammenstellung. (*Versuche* 17-18)

Following these comments, Mendel provides the reader with a chart in which his expression for the distribution of traits is applied hypothetically over multiple generations.

Generation	<i>A</i>	<i>Aa</i>	<i>a</i>	Ratios <i>A</i> : <i>Aa</i> : <i>a</i>
1	1	2	1	1 : 2 : 1
2	6	4	6	3 : 2 : 3
3	28	8	28	7 : 2 : 7
.				.
.				.
.				.
<i>n</i>				$2^n - 1$: 2 : $2^n - 1$ (14)

As a persuasive device, this chart has two functions. On the one hand, it serves as a means of gaining the support of Mendel’s readers who might be skeptical regarding the law-like nature of his findings by showing them that his ratios over time support one of their most cherished beliefs: that most plants tend to revert to parental forms. The result of this application reveals that over each successive generation the pure parental forms do increase dramatically while the hybrid forms remain constant. This suggests, mathematically, that the expression that Mendel is arguing in favor of predicts the widely supported observation of reversion.

On the other, the chart facilitates the reader’s acceptance of the law-like nature of his findings by logically transitioning them from static empirical data to dynamic mathematical relationship that the data suggest. The chart begins with the fixed ratio that he has uncovered through empirical observation of peas. It then moves the reader through successive hypothetical generations to reveal a dynamic but regular pattern of change.

This movement eventually leads them to a general mathematical formula which, if they have accepted the regular pattern of trait distribution in the empirical data and its regular action on successive generations, they are compelled to accept. This final formula— $2^n - 1 : 2 : 2^n - 1$ —is described by Mendel as *the law of development* (Entwicklungs-Gesetz).

With a mathematically describable law stipulating the regularity of the appearance of a single pair of traits over many generations in hand, Mendel moves on to prove that this law applies to a whole suite of traits in an individual organism, so long as those traits breed true before crossing. In the next section of the text, “The Offspring of Hybrids in Which Several Differentiating Characters are Associated” (Die Nachkommen der Hybriden, in welchen mehrere differirende Merkmale verbunden sind) he accomplishes this task by relying on many of the same tactics that he used to persuade his audience in the previous section. Again, he uses accumulated empirical evidence from controlled hybridization experiments to make his case and moves his audience from the static empirical results towards dynamic abstract mathematical manipulations which offer deductive support for his position.

He begins his argument in the section with natural language descriptors but quickly moves to reduce them to capital and lower case letters. This reduction allows his readers to understand what traits he is discussing without his having to write out those traits every time he discusses them.

<p>Expt. 1 – <i>AB</i>, seed parents; <i>A</i>, form round; <i>B</i>, albumen yellow.</p>	<p><i>ab</i>, pollen parents; <i>a</i>, form wrinkled; <i>b</i>, albumen green.⁴⁹ (15)</p>
---	---

⁴⁹ Erster Versuch: *AB* samenpflanze,
A Gestalt rund,
B Albumen gelb,

ab Pollenpflanze,
a Gestalt kantig,
b Albumen grün. (*Versuche* 19)

After he has established his system of notation, Mendel gradually moves from natural language descriptors of his data towards quasi-mathematical expressions to describe the combinatorial relationships between the suites of traits he discusses. This progression is apparent in the manner in which the results of the second-, third-, and fourth-round crosses are described with increasing abstraction. The first round of dihybrid self-crosses *F1* are described using numbers and natural language phrases in a fashion similar to the descriptions of the data in the first three sections of the results portion of the text, except that they are listed rather than presented within the syntax of a natural language sentence.

315 round and yellow,
 101 wrinkled and yellow,
 108 round and green,
 32 wrinkled and green.⁵⁰ (Mendel 15)

In the next section, describing the second round of dihybrid self-crosses *F2*, Mendel introduces abstraction by including the letter symbols for the traits alongside the natural language descriptions.

38 had round yellow seeds*AB*
 65 round yellow and green seeds *ABb*
 60 round yellow and wrinkled yellow seeds*AaB*
 138 round yellow and green, wrinkled yellow and green seeds ...*AaBb*⁵¹ (16)

⁵⁰ 315 rund und gelb.
 101 kantig und gelb,
 108 rund und grün,
 32 kantig und grün. (*Versuche* 19)

Following this initial chart, however, he offers a more orderly accounting of the results of the second round of dihybrid crosses without natural language.

38	plants with the sign ⁵²	AB
35	“ “ “ “	Ab
28	“ “ “ “	aB
.		
.		
.		
(16)		

As the number of possible results increases, the amount of natural language text decreases. This transition is warranted because it permits Mendel to control the clarity of his results as they increase in complexity. At the same time, however, the movement away from natural language affords him the opportunity to lead his readers from the concrete results of his experiments towards the abstract quasi-mathematical expressions that he wishes to establish as general descriptors for the pattern of inheritance. In fact, by the end of his description of the results of the dihybrid crosses, Mendel has reduced all of the relationships into quasi-mathematical expressions summarizing the result of the crosses.

⁵¹ 38 runde gelbe Samen AB
 65 runde gelbe und grün Samen ABb
 60 runde gelbe und kantig gelbe Samen AaB
 138 runde gelbe und grüne, kantige gelbe und grüne Samen $AaBb$ (*Versuche* 19)

⁵² 38 Pflanzen mit der Bezeichnung (*Versuche* 20)

Consequently the offspring of the hybrids [F2]—if two kinds of differentiating characters are combined therein—are represented by the expression [:]⁵³

$$AB+ab+Ab+aB+2AaB+2Aab+2ABb+2aBb+4AaBb. \text{ (Mendel 17)}$$

It is not hard to imagine that this neat expression of the final result of the dihybrid crosses would be cumbersome for readers if it were written in natural language. Mendel's expression, on the other hand, does a great deal to make the results accessible while at the same time leading readers towards two important conclusions. First, by transitioning through a series of forms moving from raw data expressed in natural language towards increasingly abstract quasi-mathematical representation, Mendel encourages his reader to see in the concrete semi-variable empirical data the expression of a fixed, regular natural law. Second, Mendel proves for a second time the validity of his first law stated in the expression $A+2Aa+a$ by deriving it again deductively from the expression

$$AB+ab+Ab+aB+2AaB+2Aab+2ABb+2aBb+4AaBb.$$

This expression [$AB+ab+Ab+aB+2AaB+2Aab+2ABb+2aBb+4AaBb$] is indisputably a combination of series in which the two expressions for the characters A and a , B and b are combined. We arrive at the full number of classes of the series by the combination of the expressions:⁵⁴

⁵³ Daher entwickeln sich die Nachkommen der Hybriden, wenn in denselben zweierlei differirende Merkmale verbunden sind, nach dem Ausdrücke. (*Versuche 20*)

⁵⁴ Diese Entwicklungsreihe ist unbestritten eine Combinationsreihe, in welcher die beiden Entwicklungsreihen für die Merkmale A und a , B und b gliedweise verbunden sind. Man erhält die Glieder der Reihe vollzählig durch die Combinirung der Ausdrücke. (*Versuche 21*)

$$A+2Aa+a$$

$$B+2Bb+b. \quad (17)$$

In the last portion of this section, Mendel follows the same procedures in order to prove that character traits maintain a constant ratio of distribution in trihybrid crosses. He begins again by presenting the raw data and then transitions towards a more abstract presentation as the data becomes more copious. Finally, he presents the data in the form of a quasi-mathematical expression from which he derives once again the basic formulas for the ratio of traits for a given series of crosses.

After accumulating the evidence from both the dihybrid and trihybrid crosses and showing that the ratio of characters can be mathematically deduced from their results, Mendel, drawing on his knowledge of combinatorics, concludes:

There is therefore no doubt that for the whole of the characters involved in the experiments the principle applies that *the offspring of hybrids in which severally essentially different characters are combined exhibit the terms of a series of combinations, in which the developmental series for each pair of differentiating characters are united.*⁵⁵ (19)

Now that Mendel has established through the accumulation of inductive proof and through mathematical deduction that the result of the combination of different traits in dihybrid and trihybrid crosses is always simply the combination of the separate expected

⁵⁵ Es unterliegt daher keinem Zweifel, dass für sämtliche in die Versuche aufgenommenen Merkmale der Satz Giltigkeit habe: die Nachkomme der Hybriden, in welchen mehrere wesentlich verschiedene Merkmale vereinigt sind, stellen die Glieder einer Combinationsreihe vor, in welchen die Entwicklungsreihen für je zwei differende Merkmale verbunden sind. (*Versuche 22*)

ratios for each of the individual characters, he concludes that, “*the relation of each pair of different characters in hybrid union is independent of the other differences in the two original parental stocks*”⁵⁶ (19). With this statement, Mendel issues his formal challenge to traditional theories of inheritance that emphasized a homogeneous species essence by arguing, based on his ability to mathematically extract each of the ratios independently from the combined expression of traits, that character pairs are separate and non-interacting.

Having proved deductively that characters assort independently, Mendel uses the mathematical law of combinations to predict the number of non-variable traits that should appear if all the possible combinations are made. He begins by providing the formulae for calculating 1) the number of combinations possible for a given number of character pairs, 2) the number of individuals that will occur for all possible combinations, and 3) the number of trait unions which will remain constant in subsequent generations of self-crossing.⁵⁷ Collectively, Mendel refers to these formulae as *The Law of Combination of Different Characters*.⁵⁸

If n represents the number of the differentiating characters in the two original stocks, 3^n gives the number of terms of the combination series, 4^n the number of

⁵⁶ Verhalten je zweier differirender Merkmale in hybrider Verbindung unadhängig ist von den anderweitigen unterschieden an den beiden Stammpflanzen. (*Versuche 22*)

⁵⁷ If, for example, there are two different characters pairs Aa and Bb ($n=2$) then there should be nine (3^2) possible combination possibilities $AB+ab+Ab+aB+2AaB+2Aab+2ABb+2aBb+4AaBb$ which describe sixteen different individuals (4^2) and have four (2^2) constant unions AB , ab , Ab , and aB .

⁵⁸ Das Gesetz der Combinirung der differirenden Merkmale. (*Versuche 32*)

individuals which belong to these series, and 2^n the number of unions which remain constant.⁵⁹ (Mendel 19)

Based on the laws of combinatorics, the number of forms which remain constant can be calculated by plugging the number of pairs of traits (n) into the expression 2^n . Although he does not supply the data of all of the possible combinations for seven crosses, Mendel argues that his crossings have yielded 128 constant unions, the exact number predicted by his law for seven character pairs. This fact, he argues, gives practical proof, “*that the constant characters which appear in the several varieties of a group of plants may be obtained in all the associations which are possible according to the laws of combination, by means of repeated artificial fertilization*”⁶⁰ (19). In other words, no matter what trait or variety is crossed, a predictable number of forms will always remain constant. Thus, Mendel is able to deduce and to provide empirical evidence to support his primary claim for the fixity of species.

After proving both inductively from empirical results and deductively using mathematical manipulations and principles 1) that a single trait varies regularly over many generations, 2) that multiple traits combined in a single organism assort independently, and 3) that the constant characters will always predictably reappear in exact numbers, Mendel moves the focus of his investigation to the reproductive process. In the section “The Reproductive Cells of the Hybrids” (Die Befruchtungs-Zellen der

⁵⁹ Bezeichnet n die Anzahl der charakteristischen Unterschieden an den beiden Stammpflanzen, so gibt 3^n die Gliederzahl der Combinationsreihe, 4^n die Anzahl der Individuen, welchen die Reihe gehören, und 2^n die Zahl der Verbindungen, welche constant bleiben. (*Versuche* 22-23)

⁶⁰ Dass constante Merkmale, welche an verschiedenen Formen einer Pflanzsippe vorkommen, auf dem Wege der weiderholten künstlichen Befruchtung in alle Verbindungen treten können, welche nach den Regeln der Combination möglich sind. (*Versuche* 23)

Hybriden) he provides his readers with evidence and arguments to support the conclusion that the egg and pollen cells, “represent in equal numbers all constant forms which result from the combination of the characters united in fertilization”⁶¹ (25). In order to persuade his audience to accept this conclusion, Mendel makes a classic hypothetico-deductive argument.

The hypothetical experiment he proposes to confirm the equality of both the quantity and quality of parental contributions is similar to his dihybrid experiment in the sense that he crosses two sets of plants—one with pure dominant traits for seed shape and albumen *AB* and the other with pure recessive versions of those traits *ab*. In this experiment, however, he carefully makes the point to cross exactly half of the flowers on one set of plants in the experiment with the pollen from the other half and vice versa. If the contribution of characters by each parent is equal, as he predicts, then the first round of crosses *C1* should result in offspring whose pollen and egg cells contain the following trait combinations: *AB*, *Ab*, *aB*, and *ab*. This result, of course, reflects and, therefore, reaffirms his previously posited ratio of 1:2:1.

Once he has created a set of organisms with these mixed traits in their cells, Mendel explains he will then cross the mixes with pure bred *AB* and *ab* again.⁶² He explains that if his previous assumptions are correct—that both the male pollen and the female eggs will contain the exact same qualitative features in the exact same amount—then the resulting offspring should divide out into two distinct sets of plants. In the first

⁶¹ Welche ihrer Beschaffenheit nach in gleicher Anzahl allen constanten Formen entsprechen, welche aus der Combinirung der durch Befruchtung vereinigten Merkmale hervorgehen. (*Versuche* 29)

⁶² In this second round of crosses *C2*, he will combine in equal amounts the pollen and egg cells of the *C1* mixed-trait plants with pollen and egg cells of pure-bred *AB* plants. The same procedure will also be applied in the crossing of the egg and pollen cells of *C1s* and pure-bred *ab* plants.

set,⁶³ all of the plants should express only the dominant traits AB , ABb , AaB , $AaBb$ (Mendel 21). In the second set,⁶⁴ he predicts that the offspring should present the 1:2:1 ratio of dominant and recessive traits $AaBb$, Aab , aBb , ab (21). Once he has made his predictions, he reveals to his readers the results of his actual experiments in which the data supports his experimental predictions.

The brilliance of these experiments is that they are able to provide evidence for both of Mendel's conclusions in a single set of trials. With the third set of crosses between plants with pure bred traits AB and mixed hybrids, Mendel proves that both the egg cells and the pollen cells can contribute exactly the same qualitative traits. If they had not, then half of the offspring in the resulting generation would have had dominant traits and the other half recessive. Because they all expressed the dominant traits, each of the parents had to have contributed equally to the offspring. With the second set of crossings between plants with recessive traits ab and mixed hybrids, he proves that not only are the traits qualitatively similar in the egg and pollen cells, but also that they exist in approximately the same amount on each. If they had not, then the ratios between pure bred and hybrid offspring would not have been maintained.

Unlike his previous arguments, Mendel begins making his case here by positing a hypothetical scenario using his previously established conclusions. He predicts completely new conclusions which he then verifies using experimental results. This change of logical approach shifts the emphasis towards his theoretical statements and

⁶³ In which mixed trait CI s are crossed with the pollen of a dominant AB plant and AB plants are crossed with the pollen of a mixed trait CI s

⁶⁴ Which introduces the pollen of a recessive ab plant to mixed trait CI eggs and the pollen of mixed trait CI s with the eggs of an ab

emphasizes their strength in making predictions about the natural world which, in turn, provides support for his contention that he has discovered laws of heredity.

After establishing credibility with his readers that he is describing a law of nature by making successful predictions, Mendel provides them in the final section with a complete description of the hereditary process, from fertilization to expression, based on these laws. He begins by appealing to the principles of probability to make the case that the results of self-crossing a hybrid plant should lead to a situation in which every possible combination of the dominant and recessive traits is expressed because each of these characters is equally represented in the pollen and egg cells. This argument from the principles of probability is important because it supports both the veracity of his character ratios as well as his claims of equal qualitative and quantitative contributions of characters in fertilization.

It remains... purely a matter of chance which of the two sorts of pollen will become united with each separate egg cell. According, however, to the law of probability, it will always happen, on the average of many cases, that each pollen form, A and a , will unite equally often with each egg cell form A and a ⁶⁵. (Mendel 25)

What Mendel is describing here under the title “the law of probability” (Regeln der Wahrscheinlichkeit) is actually a combination of principles. The first is the principle

⁶⁵ Es bleibt ganz dem Zufalle überlassen, welche von den beiden Pollenarten sich mit jeder einzelnen Keimzelle verbindet. Indessen wird es nach den Regeln der Wahrscheinlichkeit im Durchschnitt vieler Fälle immer geschehen, dass sich jede Pollenform A und a gleich oft mit jeder Keimzellform A und a . (*Versuche* 29)

of elementary outcomes which suggests that given a single fair and independent trial, the result of any given outcome will be $1/\text{number of possible outcomes}$. Since there are two possible outcomes in this situation, a dominant trait A or recessive trait a , there is exactly always a .5 percent chance that a particular variable will appear. Because, however, Mendel is interested not only in the individual probability of the appearance of one trait or another but in the probability of the appearance of either in conjunction with another, it is also necessary to calculate the probability of the two independent equiprobable events occurring together. This can be easily done by referring to Pascal's *Arithmetical Triangle*, which permits the calculation of the binomial coefficient or the possible successes of a particular outcome given a certain number of trials n .

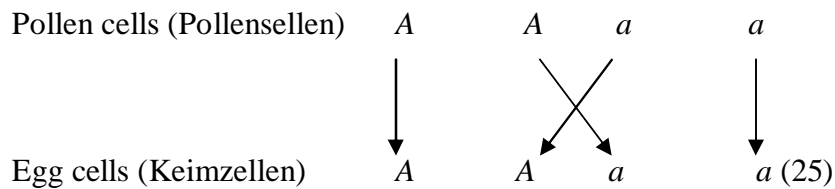
In Mendel's model, we must imagine that each joining of a dominant with a recessive trait involves the same probability as the flipping of two separate coins with the possible pairings either both dominant A , both recessive a , or one A and one a . Accepting that the probability P of each event is ideally .5 and that the number of trials n taking place in each pairing is 2, Pascal's triangle reveals that the probability distribution is $AA = .25$, $aa = .25$, and Aa or $aA = .50$ (Gonick and Smith 77). These calculations reveal that the average distribution of these traits in any given generation should be ideally 1:1:2, the exact ratio that Mendel describes in his expression $A+2Aa+a$.

Starting with the assumption that the action of fertilization is random and independent in each case, as in a probabilistic trial, Mendel presents his reader with a step-by-step walkthrough of the process of trait combination. He presents this information visually, interspersing mathematical symbols into the description and giving it the air of a mathematical proof.

He begins by describing the traits in the soon-to-be-crossed parents as equally manifested in the egg and the pollen of each. His choice of the ordering and juxtaposition of traits reinforces his position on the similarity of their contributions.

The pollen cells (die Pollensellen) $A+A+a+a$
 The egg cells (die Keimzellen) $A+A+a+a$ (25)

Once he has defined the starting scene of the process, he then proceeds to a visual diagram describing the action of crossing using arrows to indicate the direction of the coupling of the characters. This visual is, of course, based on the ideal probable outcome of the situation, rather than what might really happen in any given crossing of two characters.



Once Mendel is finished showing the action of cross-pollination combining two traits into a single plant, he symbolically presents his readers with the results of the combinations.

$\underline{A} + \underline{A} + \underline{a} + \underline{a}$
 $A \quad a \quad A \quad a$ (Mendel 25)

In this expression of the results, the symbols for division and addition play important roles. The horizontal line reinforces for the reader the idea that the traits in a new organism are both together but at the same time existing as separate, particulate, and non-interacting entities. The addition symbol “+” between the sets of traits signals that the traits are part of a series of results obtained from the same procedure. In neither of these cases do the symbols actually entail the usual mathematical relations for which they are used; however, in addition to the functions mentioned above, they suggest that Mendel is describing the steps of the process in terms of the steps of a mathematical proof.

That these symbols represent separation and membership is fairly evident in Mendel’s final transformation in which he restates the relationship of these traits in the set in terms of his familiar 1:2:1 ratio expression.

$$\frac{A}{A} + \frac{A}{a} + \frac{a}{A} + \frac{a}{a} = A + 2Aa + a \quad (26).$$

By starting with what he believes are established premises from his previous proof and from established principles in the field of probability, Mendel ends this major line of argumentation in the results section with a step-by-step deduction of the foundational ratio describing the distribution of traits in a given generation that he had derived from his experimental data in the opening sections of the paper. This allows Mendel to visually represent for his readers all of the pieces of the logical puzzle he is constructing and the mathematical neatness with which they all fit together. In doing so, he suggests the logical deductive veracity of his conclusion.

Whereas the first portion of Mendel's "Experiments" is dedicated to making the case for a regular pattern in the expression of inherited features, the focus of the second portion of the paper is on proving that these patterns are expressions of regular and predictable natural processes which, because they are regular and predictable and find expression and support in existing mathematical principles, can rightly be considered laws of nature. In the first section, Mendel argues for a mathematical law of development from the assumptions 1) that the ratios described in the first section are accurate and 2) that they hold true over every generation. In the next, he makes the case for the law of segregation by recovering his ratios for all the characters combined in bihybrid and trihybrid crosses. In the third section, he proves the independent and equal contribution of characters from the egg and sperm through hypothetico-deductive experiment: first, by predicting the outcome of his experiments based on his previously established law of development; then, by reporting the experimental results proving his predictions were correct.

In all of these arguments, mathematical symbols, operations, and principles play an increasingly crucial role. In the case of the law of development, a mathematical expression for constant unions 2^n from the laws of combinatorics is used to present the value for constant forms. This suggests that Mendel looked to combinatorics not only as a source for the descriptive notation for his theory, but also for the invention of his arguments about what the ratios of dominant, recessive, and mixed should be.

In the second argument for segregation, Mendel relies on quasi-mathematical notation and the law of combination of different characters to make the case not only that the combination and segregation of trait pairs do not influence the combination and

segregation of other traits, but also that traits breed true in a regular fashion over time. By using mathematical principles and quasi-mathematical notations and operations, Mendel lends a deductive ethos to his conclusions to convince the readers of the law-like property of his findings.

In the final section, Mendel supports his conclusion, that the traits of the male and female are equally represented and independent in the process of fertilization, by invoking the special topics of probability. In addition, he also relies on a quasi-mathematical format to suggest to his reader that his description of the reproductive process is analogous to a deductive mathematical proof of the law of development.

As a whole, Mendel's mathematization of his arguments and his use of mathematical principles to support his work make the case that characters do not vary over time. In order for phenomena to be subject to mathematical description and prediction, they need to display a regular pattern of behavior. By providing mathematical descriptions which lead to predictions, i.e. mathematical laws, Mendel boldly asserts the persistence of characters over generations.

Critical Responses to Mendel's "Experiments in Plant Hybridization"

In addition to the role that mathematics plays in making Mendel's biological arguments, the most fascinating rhetorical aspect of the story of Mendel's work is why, given its current importance to the development of the modern theory of genetics, it received so little attention when it was published. The route to publication for Mendel's work was the most common one for amateur naturalists at the time, through the

proceedings of a local natural history/philosophy society to which they belonged. In Mendel's case, this was the Natural History Society of Brunn.

The results of his experiments were first presented in two one-hour lectures in February and March of 1865 at the Society's meetings. The reaction that Mendel got from his colleagues foreshadows to some degree the reception that his paper would get from the wider scientific audience after its publication in the Society's *Proceedings*. According to Mendel's own account of the paper's reception, his findings were controversial because of their unorthodoxy and none of the Society member's felt they were sufficiently important to replicate.

I knew that my results were not easy to harmonize with contemporary science, and that in view of this publication of an isolated experiment might be doubly dangerous.... I did my best to institute control experiments, and for that reason at the meeting of the Society for the Study of Natural Science [Iltis' translators version of the society's title "naturforschenden Vereines in Brunn"] I described my experiments with *Pisum*. As was only to be expected, I encountered very various views, but, so far as I know, no one undertook a repetition of the experiments (Iltis 180).

Even though the Society's members found Mendel's results controversial at best and forgettable at worst, his complete lecture was published, the customary procedure for all papers presented at the Society's meetings, in the *Proceedings of the Brunn Natural History Society* (Henig 142). Undaunted by the lukewarm response he received from the

Brünn Society, Mendel requested forty reprints from the journal editor so that he might share his findings with important figures in hybridization and related fields in hopes that they might recognize and support what he believed were extremely important results.

Of the forty reprints Mendel requested, twelve are known to have been sent out. Among the recipients are some of the most important figures studying variation, evolution, cytology, botany, hybridization, and reproduction at the time Mendel was writing. The most famous recipient was Charles Darwin, who was found to have a reprint on his shelf, the pages uncut. Perhaps Mendel's motivation for sending Darwin the paper was a belief that his laws were relevant to Darwin's struggle to understand the hereditary process and its role in variation. Perhaps he believed that his theory of non-variable characters, though on the face of it contradictory to the Darwinian doctrine of continuous variation, might help Darwin solve the problem of the dispersion of useful variation over many generations of breeding. Why Darwin didn't open the reprint is unknown. Perhaps it was because it was from a relatively obscure journal put out by a small natural history society. Perhaps it was because he had never heard of Gregor Mendel before.

Though a language barrier might have contributed to Darwin's disinclination to look at the paper, this same factor cannot be said to have influenced the prominent German and Austrian, and perhaps even the Dutch scientists who received copies. The recipients included scientific luminaries such as Martinus Beijerinck, a Dutch biologist and co-discoverer of viruses; M.J. Schleiden, the establisher of cell theory; and Carl von Nägeli, a well-respected cytologist who worked with Schleiden. A fifth copy was sent to a classmate of Mendel's from Vienna, Kerner von Marilaun, who was at the time of the manuscript's publication a biologist in Innsbruck. There is also speculation, but no

substantive proof, that a sixth reprint was sent to Franz Unger, Mendel's and Marilaun's botany professor from Vienna. Although six other copies are known to exist, there is no evidence of who their original owners were (Henig 142-46).

Among the German and Austrian biologists and cytologists there is a common thread of connection which suggests why Mendel believed they would be receptive audiences for his mathematically described theory of inheritance. During the 1840s and 1850s in Germany and Austria, there was a small but persistent trend in research away from the qualitative descriptive methodologies of traditional hybridists and towards the more disciplined experimental/mathematical approach used in the physical sciences. According to Robin Henig, this method, described in the work of the cytologists Matthias Schleiden and Theodor Schwann, gained traction under the moniker "scientific botany" and found supporters in both Nägeli and Unger (Henig 57). Given the important role of mathematics in Mendel's arguments, it seems only fitting that he would target a group of biologists who would appreciate his attempt to express his conclusions mathematically and use mathematic principles to make his arguments.

Of this group of influential and mathematically inclined biological researchers, Carl von Nägeli was the only known to have fully read, considered, and responded to Mendel's work. His critique of Mendel's theory offers us the only solid evidence about what concerns the recipients who did read the paper might have had with it.

The reason that Mendel sent Nägeli his theory appears to be his belief that the cytologist would appreciate and support his mathematical approach to heredity. However, there is evidence that Mendel was not certain that the common ground of methodology was sufficient to solicit Nägeli's attention to his work. To add relevance to his work for

the cytologist, he also suggested in the letter which accompanied the monograph that he was interested in doing further hybridization experiments on *Hieracium* (Hawk Weed) a species on which Nägeli was an expert.

The correspondence between the two suggests that Mendel's expressed interest in the *Hieracium* experiments was an important factor in Nägeli's motivation to write back. In his response to Mendel's first letter and his article, the only existing communication of Nägeli's to Mendel, Nägeli writes at length about *Hieracium* and asks for Mendel's help in doing some breeding experiments with it.

Your design to experiment on plants of other kinds is excellent, and I am convinced that with these different forms you will get notably different results (in respect of the inherited characters). It would seem to me especially valuable if you were able to effect hybrid fertilizations in *Hieracium*. (Iltis 192)

In this response, Nägeli provides evidence not only of his interest in Mendel's proposed experiments with *Hieracium*, but also of the reasons why he believes the results of Mendel's first experiments are aberrant if not outright wrong. First, his statement that Mendel will "get notably different results" suggests that he rejects Mendel's primary conclusion that character traits remained unaltered and regularly distributed through the process of combination and separation. Nägeli's rejection of this position is not surprising given that he is a firm believer that during fertilization the characters of the parents typically blend to create a new intermediate character. He expresses these views in an

article “Hybridization in the Plant Kingdom⁶⁶” (Die Bastardbildung im Pflanzenreich) published in the very same year as Mendel’s “Experiments.”

The rule, however, is that the characters of the father and the mother combine and interpenetrate, whereby a new individual character originates which holds more or less the mean. The way and the manner in which the union occurs cannot be determined in advance. (Quoted in Roberts 96)

In addition to believing that Mendel’s conclusion was incorrect, Nägeli also provides his opinion about why Mendel got it wrong. In his estimation, Mendel fails to make his case because he has not done a sufficient number of experiments. Although this seems like a ridiculous accusation given that Mendel’s conclusions were based on the crosses of nearly 10,000 pea plants, Nägeli’s complaints are not that Mendel didn’t do enough experiments with peas but that he didn’t do a sufficient number of experiments with other species. Because his sample is deep but not wide, Nägeli argues that Mendel cannot assume that his belief in the inalterability of traits is good in all cases.

Nägeli’s comment about the breadth of Mendel’s sample may have also encouraged him to attack Mendel’s move in his argument to present his empirical findings as the source for a predictable, mathematically describable, law of nature. According to Mendel, in his second letter to Nägeli written April 18, 1867, Nägeli had written him that he, “should regard the numerical expressions as being only empirical, because they cannot be proved rational” (Stern and Sherwood 63). The point that Nägeli was trying to make to Mendel with this piece of criticism is that Mendel could not rely on

⁶⁶ Karl von Nägeli, “Die Bastardbildung im Pflanzenreich,” *Botanische Mittheilungen* 2 (1865): 187-235.

rational proof from the principles of mathematics to support his argument for the regularity in the distribution of traits in hybridized peas. In his reply, Mendel defends his use of deductive reasoning to argue his position.

My experiments with single traits all lead to the same result: that from the seeds of hybrids, plants are obtained half of which in turn carry the hybrid trait (Aa), the other half, however, receive the parental traits A and a in equal amounts....

Therefore $2Aa+A+a$ or $A+2Aa+a$ is the empirical simple series for two differing traits. Likewise it was shown in an empirical manner that, if two or three differing traits are combined in the hybrid, the series is a combination of two or three simple series. Up to this point I don't believe I can be accused of having left the realm of experimentation. If then I extend this combination of simple series to any number of differences between the two parental plants, I have indeed entered the rational domain. This seems permissible, however, because I have proved by previous experiments that the development of a pair of differing traits proceeds independently of any other differences. (Stern and Sherwood 63)

In his reply, Mendel argues that because he has shown that the same pattern describable with the same mathematical expression can be found in independent empirical observations across at least three different traits, he should be able to assume that it will be the same for all traits through subsequent generations. Though Nägeli is no stranger to the application of mathematics to describe biological phenomena in his own

work,⁶⁷ he was not particularly enthused about Mendel's use of mathematical claims to support a theory which directly opposed his own position, especially because Mendel's ability to mathematize his claims served as a compelling argument that Mendel's belief in the non-variability of characters should be accepted.

By investigating Nägeli's response to Mendel and Mendel's defense of his own work, it is possible to understand at least in Nägeli's case why Mendel's work was not compelling. To some degree it failed to make a strong impression because it supported the conservative theory held by Kölreuter and Gärtner that the essence of the species remained separate and inalterable in the offspring of hybrids and tended to separate out over multiple generations. This position was in conflict with Nägeli's position that a completely new character was formed from the blending of the parents' essence during the process of hybridization.

The concept of non-variable traits was also in conflict with evolutionary theories, both Darwinian and Lamarkian, which had gained ascendancy during the period Mendel was distributing his reprints. Because many biologists including Unger and Nägeli had embraced this position, it is likely that at least some portion of Mendel's audience rejected his findings on evolutionary grounds.

In addition to challenging the substance of Mendel's conclusions, the belief that organisms vary from one generation to the next or blend to form new characters also challenged the manner in which Mendel argued for, arrived at, and presented his conclusions. If organisms vary and or new characteristics are formed in each generation, then the laws of combinatorics and probability which Mendel used to design his

⁶⁷ Iltis explains that in Nägeli's cytological work his research, "led him to... formulate upon a mathematical basis his theory of growth [in which].... Cells constitute the elements out of which we can construct the organs in accordance with mathematical rules" (Iltis 184).

experiments and to deduce his laws are inapplicable because these laws require outcomes to be as fixed and clearly discernable as the results of a coin toss. If characters vary or blend, neither of these conditions is possible.

Though mathematical models developed in the 1920s and 1930s and the recognition that mutation was a natural source of variation allowed for a grand synthesis between Mendel and Darwin's work, at the time Mendel was sending out his work it would have been impossible for his audience to reconcile his conclusion with Darwin's. Given the general excitement during that period generated by *The Origin*, it is not surprising that Mendel's insistence on the persistence of seven traits in a single species of plant did not draw much attention. In fact, it was not until Darwin's concept of continuous variation was repeatedly challenged and Mendel's belief in the stability of characters was three times confirmed by De Vries, Tschermak, and Correns that his theory was heard by a more receptive audience in the beginning of the twentieth century.

Conclusion

An investigation of Mendel's rhetorical situation, his background, and his arguments reveals the complexity of forces affecting the rhetorical situation surrounding the publication of "Experiments in Plant Hybridization." As a participant in the discourse regarding plant hybridization, Mendel was guided by the information, methodologies, and arguments that had come before him. His work shows knowledge of and reverence for the observations of the hybridists who preceded him, though his methods for confirming their position on the fixity of character traits were radically different.

Drawing from his experiences as a student of mathematics and the physical sciences, Mendel applied the hypothetico-deductive approach to his experiments which required him first to state his proposed conclusions and then plan experiments to prove their validity. In this planning process he relied on the principles of combinatorics and probability to supply him with direction in the a priori framing of his investigation. By starting with these principles to develop his hypothesis, designing an appropriate set of experiments, and rigorously carrying out those experiments, Mendel was able to demonstrate the regularity of the distribution of characters over generations, the independent segregation and integration of character pairs, and the independent and equal contributions of parental characters in offspring.

Mendel is interesting rhetorically, on the one hand, because his arguments represent a novel first case in which mathematics begins to be employed to make arguments about heredity. A close examination of his “Experiments in Plant Hybridization” reveals that Mendel used methods of deductive reasoning to elevate his empirical results to the status of a natural law. In order to persuade his readers that traits are not variable across generations and that their appearance follows a law of nature, Mendel used quantified data, argued for regularity in the data using mathematical and quasi-mathematical formulae, and elevated these patterns to the status of a law through the use of quasi-deductive proofs and the argument that the regular patterns he described follow existing laws of combinatorics and probability.

On the other hand, Mendel’s failure to find an audience for his work, despite the eventual importance of his conclusions to modern genetics, represents something of a modern rhetorical mystery. How could such a clearly argued and currently celebrated

paper have been ignored for so many years? A detailed examination of his rhetorical situation, background, and arguments, however, reveals that the reasons why Mendel's work was ignored are numerous. Perhaps his work was passed over because of language barriers. Perhaps it was overlooked because the journal he published in wasn't held in the highest esteem. It could have been ignored because Mendel was a geographically and academically obscure figure.

Though Nägeli's response to his work gives no hints about whether these possibilities were a factor, it does provide evidence that he was overlooked because his work challenged the theory of evolution which expected variation in each generation. In addition, it made great claims about inheritance based on experiments on a single species. Without social or professional credibility, which might encourage his readers to look past these issues, it is perhaps not surprising that Mendel went unrecognized for so many years.

Chapter 4: Contours of Heredity: The Law of Error and Francis Galton's Biometrical Approach to Heredity

*To those who are acquainted with the statistics set forth by Galton in *Natural Inheritance*...it will be well known that measurements of certain quantities...group themselves around a mean form in such a way that the curve representing the frequency of occurrence of the several measurements has the form known as a curve of Frequency of Error. – William Bateson in *Variation in Secondary Sexual Characters*⁶⁸*

It was Galton who first freed me from the prejudice that sound mathematics could only be applied to natural phenomena under the category of causation. Here for the first time was the possibility... of reaching knowledge—as valid as physical knowledge was then thought to be—in the field of living forms and above all in the field of human conduct. – Karl Pearson⁶⁹

Although it may seem historically that Mendel's "Experiments in Plant Hybridization" represents a lone voice advocating a systematic mathematical approach to the study of heredity informed by the special topics of probability, the reality is that he was not the only researcher in the latter half of the nineteenth century applying mathematical principles to the investigation of inheritance. In fact, a decade before Mendel's arguments on the subject were widely read and appreciated in England, Sir Francis Galton (1822-1911) broke ground for a probabilistic/statistical approach to the investigation of heredity in his book *Natural Inheritance* (1889).

Like Mendel, Galton develops in his work a theory of heredity which employs quantitative evidence and the special topics of probability. However, while Mendel used basic principles of probability codified as combinatorics, Galton used the law of error to establish meaningful patterns in hereditary data. Also, Galton's efforts to persuade his readership of the efficacy of mathematical principles in describing patterns in inheritance

⁶⁸ R.C. Punnett., ed., *Scientific Papers of William Bateson*, 2 vols. (Cambridge: Cambridge University Press, 1928) 194.

⁶⁹ Nicholas Gillham, *A Life of Sir Francis Galton* (Oxford: Oxford University Press, 2001) 277.

were a success, whereas Mendel's were a failure. The rhetorical issue in this chapter, therefore, concerns what unique characteristics of the rhetor, the argument, and the audience may have resulted in the successful reception of Galton's arguments where Mendel's arguments failed. After examining each of these factors in detail, I conclude that Galton's success is due to a combination of 1) the degree of notoriety and respect he commanded in the scientific community, 2) the inclusiveness of his views about the hereditary process, and 3) the extensive rhetorical effort he made to persuade his audience that the special topics of probability and statistics have jurisdiction in making arguments about variation and its distribution in human populations. The use of mathematics to inform and describe his arguments and the acceptance by the biological community of this method of argumentation are important because the success of Galton's arguments set a precedent for employing mathematical formulae and principles in arguing about heredity in English biology, helping to establish a positive atmosphere for the reacceptance of Mendel's work.

Francis Galton: a Gentleman of Science

One of the factors in Galton's success in his arguments for a mathematically informed theory of heredity was his status in English social and intellectual circles, a status due in part to an accident of birth and in part to his devotion to science. From birth Galton enjoyed a comfortable position within the wealthiest strata of the English middle class. His father, Samuel Tertius Galton (1783-1844), was a successful Quaker

industrialist and his mother, Violetta (Darwin) Galton, was the daughter of Erasmus Darwin and aunt to Charles Darwin (Brookes xv, 4).

In addition to being born into a family of material wealth, Galton was also born into one with intellectual fame as well. His father Samuel was an amateur scientist who contributed to the theory of color vision and his grandfather Erasmus Darwin, a member of the intellectual society the Lunar Club, devised a steering mechanism for carriages, published a classic paper explaining the formation of clouds, and crafted the poem *Zoonomia* which argued that all living things were descended from a single microscopic ancestor (Bulmer 3). The propensity and talent for science of these illustrious forbearers seem to have been genetically gifted, or at least Galton would have thought so. He was considered as a child prodigious in his capacity to learn.

In the formal system of education Galton's brilliance was not so obvious. In his opinion, his lack of enthusiasm for his studies was due to the restrictiveness of the subjects he was forced to study. In his autobiography he writes of his early educational experience: "I learnt nothing and chafed at my limitations. I had craved for what was denied, namely, an abundance of good English reading, well-taught mathematics, and solid science" (*Memories* 20).

At the age of sixteen, he was given the chance to pursue his scientific interests a little more freely by his mother who hoped that Francis would join his grandfather and his uncle in the medical profession. He was sent for a year to study medicine at Birmingham General Hospital and the following year to King's College Medical School in London. In these more scientific settings, Galton's academic performance briefly improved (Bulmer 5).

In 1840, however, he left King's College to read mathematics at Cambridge. This interruption had been suggested not only by his father, but also by his cousin, Charles Darwin. Francis agreed to it whole heartedly and entered Trinity College Cambridge in October of 1840. His academic career at Cambridge, however, like his earlier secondary education, was lackluster.

About his studies in mathematics, Galton says relatively little. He admits that he had a little catching up to do in elementary mathematics when he first arrived and that he had several mathematical coaches during his time at school. In his third year reading for mathematics, however, he suffered a mental breakdown which ended his studies in the subject (*Memories* 79)

Galton's struggles to keep his restless mind confined to a rigid institutional academic program ended in 1844 with the death of his father. The considerable wealth that he inherited left him financially independent and free of his responsibility to finish medical school. He quit school, abandoned the medical profession, and took off for the Near East to travel and experience the world.

From 1845 to 1846 he traveled to Egypt and Syria, embarking on an expedition up the Nile with two Cambridge friends, Montague Boulton and Hedwith Barclay, and from there traveling to Jerusalem and Jordan. He returned to England in 1846 after facing potential trouble from the family of a guide who was killed during the journey and after the death of one of his sister's husbands.

After four years of hunting and general idleness in England, Galton decided to return to adventuring, this time in southern Africa. The importance of this exploration to Galton's interest and success in science cannot be underestimated. In the planning and

execution of the trip, Galton made a serious connection between his life's pursuits and science for the first time (*Memories* 122-23). In addition, the trip provided him with his first opportunity to gain recognition for scientific work.

The inspiration to make the trip a journey for science initially came from Francis' cousin Douglas Galton, a Fellow of the Royal Geographic Society, who suggested that Francis might incorporate in his adventure some scientific activity. He provided Galton with introductions to some of the Geographical Society's members who worked with him to develop a plan of exploration for his journey which was to take him from Cape Town to Lake Ngami, recently discovered by David Livingstone. Members of the Society suggested that Galton survey the land on his route and report back his findings (Bulmer 11). Although unversed in the use of the sextant and techniques of surveying, through reading and practicing, Galton became competent enough to use the basic tools and techniques of geodesy to survey the country through which he traveled, making careful measurements of longitude and latitude (*Memories* 125-26). As promised, he sent his observations back to the Royal Geographic Society which published them in their journal in 1852. In 1854, because of his work in Africa, he received one of their two annually awarded gold medals, "for having at his own expense and in furtherance of the expressed desire of the Society, fitted out an expedition to explore the center of South Africa, and for having so successfully conducted it... as to have enabled this Society to publish a valuable memoir and map in the last volume of the journal" (150).

Galton's success in his African adventures helped him establish a place of prestige in the scientific community and encouraged him to devote his time, energy, and wealth to the pursuit of science. Of the ethos within the scientific community that his

work in Africa afforded him Galton writes, “The Geographical Medal gave me an established position in the scientific world. In connection with subsequent work, it caused me to be elected a fellow of the Royal Society in 1856, and to receive the very high honor of election to the Athenaeum Club” (151).

With these new laurels and the support of a winning personality and disposable income to help fund scientific projects, Galton was soon an established member of the scientific elite. He was elected to the council of the Royal Geographic Society in 1854 and made secretary in 1857. He was also elected to various high positions in the British Association for the Advancement of Science including general secretary (1863-67), president of the Geographical section (1862, 1872), and president of the Anthropological section (1877, 1885), turning down the position of President of the entire society twice (Gillham 105). He was selected to be a member of the Royal Society in 1856 and in 1854 received the silver medal from the French Geological committee (105).

In addition to supplying him with the necessary credibility for prosecuting his later mathematical arguments about heredity, Galton’s early scientific exploits prepared him with the requisite mathematical skills to develop his arguments. In his African travels, his work with geodesy most likely put him in contact with some of the mathematical developments in the field which could have included work estimating error using distribution tables calculated from the error curve.

In addition to geodesy, his scientific investigations of weather may have also provided him with the necessary mathematical exposure to understand and make arguments about heredity using the error curve. In the early 1860s, Galton became interested in the laws of weather which were at the time beginning to be comprehended

with the help of the collection of numerous observations made at the same time over a wide area. Specifically, he was interested in the phenomena of wind and how it was created. At the time, the only wind model was that of the tropical cyclone in which an area of extreme low pressure and a mass of warm moist air moving in different directions collided. Galton wondered whether the direction of wind in Europe might not be caused by a similar weather phenomenon.

In order to find out, he sent out requests to observatories, lighthouses, and ships to make measurements and observations three times a day of cloud cover, wind speed and direction, temperature, and barometric pressure. He sent charts of his compiled data to the Royal Society in 1862 and in 1863 offered an analysis of the data which made the case, that not only did Europe experience cyclonic conditions similar to those in the tropics, but that it also experienced anticyclonic conditions when, during a period of high pressure, air from the upper atmosphere would rush down creating windy conditions (*Memories* 229-31).

Michael Bulmer suggests that the papers and charts compiled by Galton in his work with the anticyclone and other meteorological problems reveal his knowledge of statistics/probability. Specifically, he suggests that Galton's "facility for interpreting maps and charts, in particular by drawing isobars, isotherms, and so on through points of equal geometric pressure, temperature, and other physical qualities" seems to foretell his technique of drawing ellipses through data points on tables of the joint frequency distribution of the heights of parents and offspring to discover the properties of the normal curve in the distribution of the data. Additionally, Bulmer explains that in an 1870 paper discussing the problem of predicting wind velocity, Galton calculates wind velocity

from a crude multiple regression formula which could be seen as a precursor to the ones he develops for his hereditary experiments with peas discussed in the next section (Bulmer 31). Finally, Bulmer argues that it may have been his work with weather that put him in contact with Quetelet, who might have been his supplier for the Belgian data for the weather maps he was compiling in the winter of 1861-62 (173).

The Development of a Mathematical Theory of Inheritance

At around the same time as his meteorological investigations, Galton began work on his hereditary investigations. Even in his earliest efforts, it is possible to see the importance of quantified data in his work. At some point between 1865 and 1869, Galton was introduced to the application of the error curve through his friend William Spottiswoode, who had used the law of errors in order to provide evidence that mountain ranges were “not accidental and that the geologist and the natural philosopher will at least have good grounds for seeking some common agency which has caused their upheaval”⁷⁰ (Gillham 158).

Further, he was introduced to the physiological and sociological applications of the law through the work of Quetelet. In *Memories* Galton says that he made the acquaintance⁷¹ of the Belgian statistician (*Memories* 304). Though it is not clear what transpired during their meeting(s), after reading an English translation of his 1849 text, *Letters on Probability*, Galton decided to incorporate Quetelet’s curve fitting techniques

⁷⁰ William Spottiswoode, “On Typical Mountain Ranges: an Application of the Calculus of Probabilities to Physical Geography,” *Journal of the Royal Geographic Society* 31 (1861): 149-54.

⁷¹ Bulmer suggests that this meeting most likely took place either through his collection of weather data or at the International Statistics Conference which took place in London in 1860. (Bulmer 173)

into his assessment of his collected data on talent and its relationship to heredity. The results of this application are presented in his first extended work on heredity, *Hereditary Genius* (1869).

In this text, Galton argues for the first time that the law of errors accurately describes the distribution of variation in the traits of human populations. Given that Quetelet's statistical/probabilistic approach had yielded such wonderful results in predicting the distribution of qualities like height and chest breadth in a population, Galton argues that the same distribution might be applied to traits such as intelligence.

In the third chapter of *Hereditary Genius*, entitled "Classification of Men According to their Natural Gifts," Galton uses Quetelet's application of the law of error to a collection of heterogeneous subjects to move his reader towards his own conclusions regarding the distribution of eminence within a population.

I propose in this chapter to range men according to their natural abilities, putting them into classes separating them by equal degrees of merit, and to show the relative number of individuals included in the several classes....The method I shall employ for discovering all this, is an application of the very curious theoretical law of deviation from the average. (*Hereditary Genius* 26)

Given that Quetelet's statistical approach had yielded such wonderful results in predicting height and chest breadth, Galton suggests that the same distribution might be applied to non-physical traits such as intelligence. Thus, one of the differences between Quetelet's and Galton's approach is that it expands the application of the error curve from

easy-to-quantify physical characteristics to more difficult-to-measure qualitative features (32).

In the sections of the text which follow, Galton analyzes actual scores from the Royal Military College at Sandhurst's entrance examination using Quetelet's calculations of the law of errors and shows that the scores fall approximately within a distribution predicted by Quetelet's tables of calculations (*Hereditary Genius* 33). This proves for Galton that the distribution of intelligence, like height and chest size, can also be estimated within a population.

Although Galton's application of the law of error to the data is strictly in keeping with the applications as they are described by Quetelet, he interprets the results differently. While Quetelet is interested in ascertaining the profile of the "average man" and views deviations as errors in nature, Galton is interested in variation itself. He sees the distribution of values of a trait around a mean not as a series of mistakes in nature but as the description of the degree of variation of a particular feature in a particular natural population (26-30).

Galton explains that a natural population within a stable environment will have a predictable range of characters with a predictable "normal," i.e. bell shaped distribution. He adds that if two populations from different environments are compared, two different ranges of characters with different population means will become apparent. As a result, the law of error can be used to determine whether or not, "the events of which an average has been taken, are due to the same or to dissimilar classes of conditions" (*Hereditary Genius*, 29). Because the law of error can be used not only to describe the range and distribution of variation within a population affected by the same conditions, but also

whether a sample contains two distinct populations, it becomes a powerful tool in Galton's later work and in the work of the biometricians for defining homogeneous populations and identifying evolutionary changes within or between populations.

In addition to describing variation, Galton also occupies readers of *Hereditary Genius* with a discussion of the degree to which the intensity of a particular inherited trait is maintained across generations. To ascertain the degree, Galton examines the pedigrees of eminent men in various professions and attempts to calculate the number of eminent offspring, forbearers, and relations by blood that could be traced from a given set of eminent individuals. What he finds is "that the proportion of men of eminence continually decreased as one moved away from the most eminent person roughly according to the law that, 'the percentages are quartered at each successive remove, whether by descent or collaterally'" (287). This process of geometrical decrease eventually becomes a central feature of his argument in *Natural Inheritance* and is later codified by Galton as the *Law of Ancestral Inheritance*.

In *Hereditary Genius*, Galton develops the germs of several principles which play important roles in guiding his later arguments about the nature of heredity. First, he sets forth the principle that the distribution of a trait in any given homogeneous human population is approximately normal; and second, that the intensity of traits diminishes over each successive generation by a geometrical regression. From these principles, Galton directs his natural inquiries from specific instances of single characters to general instances of multiple characters over multiple generations in hopes of discovering 1) whether the normal distribution of traits in small homogeneous populations might also

occur in the heterogeneous aggregate of all populations, 2) how heredity affects the distribution of traits, and 3) how traits could remain stable over time (Chatterjee 288).

In order to answer these questions, Galton performs a crucial set of experiments on the same natural subject as Mendel, peas. Beginning in the spring of 1874, he designed an experiment in which he separated groups of ten peas into seven classes according to their seed weight and diameter. Galton treated these classes as homogeneous subpopulations in a larger more heterogeneous general population of peas. He sent samples from each of the seven groups to different friends to raise according to strict instructions. When the seeds were harvested, Galton carefully weighed/measured the offspring for each group, calculated the mean, and worked out the distributions of the weights and diameters of each of the offspring around the mean. What he found was that in each of the seven groups the values of their weights distributed normally around the group mean.

Now that Galton was certain that both the smaller homogeneous weight/diameter groups and the larger more heterogeneous population of a single generation distributed normally, he felt confident in comparing seed weights/diameters across generations. This would allow him 1) to ascertain any change in the expression of a particular trait from one generation to the next and 2) to determine whether the mean and the distribution of the population changed over time. In order to do these things, he invented a totally new mathematical procedure called regression analysis in which two different sets of variables, one dependent (marked with respect to the y axis) and one independent (marked with respect to the x axis), are plotted with respect to one another. Then, a line, called a regression line, is drawn which minimizes the spread in the data points. The

slope of this line provides a quantified description of whether and to what degree the dependent value of (y) can be predicted by the independent value of (x).

In his examination of two generations of peas, Galton wanted to know to what degree the magnitude of an offspring's trait (y), such as weight, could be predicted from the magnitude of the same trait in the parent (x). Plotting each of the means for each of the offspring groups (y) against each of the means for each of the parental groups (x), Galton observed that the means lay roughly on a straight line. After drawing the regression line, he measured the slope of the line to find that the predictability of the median weight of the offspring (y) from the median weight of their parents (x) was $1/3$. Galton interpreted this result as meaning that offspring had only $1/3$ of the original variability of their parents, either above or below the mean, which meant that on the whole the offspring tended to regress towards the population mean.

Given the regularity of the distribution of the means in each of the generations and the regularity of the variance of those means from the line of regression, Galton affirmed his hypotheses that 1) distributions of traits follow the principles of the law of error across generations and 2) traits regularly regressed towards the population mean. With regards to heredity, he believed that his findings showed that characters were generally stable over time, the same position supported by Mendel, because there was a constant tendency towards the population mean rather than away from it and because the traits were normally distributed over generations.

After developing his method of bivariate, i.e. two-variable, analysis on plants, Galton was eager to apply it to human subjects. His main problem, however, was that there was no available data. He decided to collect it himself. This Herculean task began

in 1884 when he sent out a notice offering a reward for statistical data from family records for such characteristics as stature, eye-color, temper, and artistic faculty (*Natural Inheritance* 72, 77). In the spring of the same year, he also set up his anthropometric laboratory at the International Health Exhibition in the Gardens of the Royal Horticultural Society in which visitors paid three pence to have their vital statistics including eyesight, hearing, height, and strength recorded (Gillham 211-12). From these statistics, in addition to the statistics from a more permanent anthropometric laboratory set up in the Scientific Galleries of the South Kensington Museum, Galton was able to obtain information on more than 17 different characteristics for over 13,000 individuals.

By the time Galton was ready to make his arguments in *Natural Inheritance*, he had already established for himself through his family connections, wealth, exploits, and hard work a significant reputation in the scientific community. Thus, Galton had access to and the attention of many influential members of the scientific community. Because of his experience with and track record of publishing on inheritance, he also had the credibility to speak authoritatively on this subject and to be taken seriously by them.

In addition to giving him credibility, his early scientific work exposed him to and gave him practical experience with statistical methods, which played a major role in developing his later work with heredity. Through the application of the error curve to biological data on humans in *Hereditary Genius* and his work with peas, Galton began to develop some of his fundamental tools for dealing with variation within and across generations of human populations. The application of these tools to plants led to results which encouraged further exploration of the phenomena of regression and the geometric diminution of traits in humans.

Natural Inheritance

Though Galton's social and scientific ethos play an important role in the acceptance of his theories, the success of *Natural Inheritance* can also be attributed in part to his skill as a rhetor. The next few sections will focus on the arguments in the text itself. The goal will be to understand not only how persuasive tactics helped secure the positive reception of the text, but also to comprehend how fundamental assumptions which ground Galton's model made it amenable to its readers.

Although up until the point that Galton published *Natural Inheritance* he had been offering various arguments about the nature of inheritance which relied in some part on the law of errors, this text represents the first time he makes an extended coherent argument for his theory of heredity based on the mathematical operations and principles developed from the law of error. After a first chapter that establishes his ethos and promises rewards for those readers who can persevere through his introduction of new mathematical methods, Galton provides his reader with a description of the hereditary process in chapters two and three which necessitates, or at the very least legitimates, the application of special topics from these two mathematical fields. Then, in chapters four and five, he makes the case that the law of errors can be applied to the study of heredity by developing his own particular interpretation of the law, making the case that this interpretation does not violate the law's basic precepts, and presenting arguments to establish that the data on hereditary features distribute in accordance with the scheme of the law of errors. In conjunction, these chapters 1) establish the basic biological principles that Galton is arguing from and 2) persuade his readership of biologists and

other scientists, who, he assumes, have either limited or no general background in the law of errors, to accept the application of this law to the study of heredity in humans.

Chapter One: The Author, the Audience, and the Journey

In the first chapter of *Natural Inheritance*, Galton supplies the reader with three very important pieces of information. First, he presents his own credentials. Second, he identifies the characteristics of the audience he imagines reading the text. Finally, he provides a brief qualitative characterization of the questions he intends to ask and the method he intends to employ to answer them.

He begins the first chapter by establishing his ethos for his reader and the reasons why his particular theory of inheritance is superior to previous theories on the same subject. In order to establish his credibility, Galton explains that he has been investigating heredity for many years and that he has published many technical memoirs concerning the subject. This appeal to past experience and success in publication lets the reader know that the theory he is about to describe is not some radically new idea, but rather one that has grown up over time: “I have long been engaged upon certain problems that lie at the base of the science of heredity, and during several years have published technical memoirs concerning them” (1).

This claim to credibility is important because Galton admits that his theory of inheritance is different, and he believes superior, to all previous theories. (He is, of course, unaware of Mendel’s work.) He explains that its novelty and strength derive from his 1) examining multitudes or populations of organisms and 2) employing, “more refined

and searching methods than those usually employed in hereditary inquiries” (1). The methods he describes are in fact the mathematical ones developed from the law of error. He proposes to his reader that by employing the mathematical operations and principles of this mathematical law he is able to provide a more precise and credible theory of heredity than those who have come before him: “The inquiry relates to the inheritance of moderately exceptional qualities by brotherhoods and multitudes rather than individuals, and it is carried on by more refined and searching methods than those usually employed in hereditary inquiries” (1).

Once he has made an ethical appeal and an argument regarding the superiority of his approach, Galton recognizes his audience and makes the case for the style and organization of his argument. One common element that he does not divulge in the introduction is his conclusions, though he does suggest a series of problems that he intends to address in the text.

The conclusions cannot however be intelligently presented in the introductory chapter. They depend on ideas that must first be well comprehended and which are now novel to the large majority of readers and unfamiliar to all. But those who care to brace themselves to a sustained effort, need not feel much regret that the road to be traveled over is indirect, and does not admit of being mapped beforehand in a way they can clearly understand. It is full of interest of its own. It familiarizes us with the measurement of variability, and with the curious laws of chance that apply to a vast diversity of subjects. (*Natural Inheritance* 3)

In these lines, Galton makes several important rhetorical moves that set the tone for the rest of the text. First, he identifies his audience. He recognizes that they are most likely unfamiliar with the operations and principles connected to the law of errors. This suggests that his argument has been designed primarily for biologists and not for those in the fields of geodesy, astronomy, mathematics, and economics, who would most likely have knowledge of the role of the error curve and would, therefore, not require an extended explanation of how the law works.

Once he has identified his audience, he tries to develop his ethos further by showing them that he understands they might have difficulties with the mathematical elements of the arguments. He assures them that their patience with the circuitous route of his arguments and their perseverance will pay off in their knowledge of a new, powerful method of investigating heredity that will lead them through the tangle of traditional problems towards a clearer understanding of the phenomena. In this exhortation, he makes the case that that which is harder is better. This argument places a premium value not only on the endurance of his reader but also on the value of the methods employed in and the results reached through his arguments.

Collectively, Galton's introductory remarks serve as a warning and a promise. They warn the reader that the approach is unorthodox and difficult. However, they promise the reader that Galton himself is sufficiently knowledgeable to lead them through this difficult approach and that if they follow and persevere, they will reap the rewards of this new method.

Chapters Two and Three: Establishing Jurisdiction

Once he has exhorted readers to follow his lead, Galton begins the task of persuading them that the mathematical methods he will employ in the text are fitting for the study of heredity by making the case that heredity is a random phenomenon and can, therefore, be best handled by the laws of probability. He begins chapter two of the text “Process in Heredity” by explicitly stating his goal.

A concise account of the chief processes in heredity will be given in this chapter, partly to serve as a reminder to those to whom the works of Darwin especially, and of other writers on the subject, are not familiar, but principally for the sake of presenting them under an aspect that best justifies the methods of investigation about to be employed. (*Natural Inheritance* 4)

In addition to stating the goal of the chapter, this passage provides further information about Galton’s intended audience for the text. Although the basic gist of Darwin’s work was widely known by 1889, the particulars of his theory were probably not common knowledge for all but a very small group of specialists in the biological sciences. Galton’s recognition here that he needs to explain the basic tenets of Darwin’s principles relevant to his work suggests that he is targeting his arguments to a broad, educated audience who might know something about the rudiments of biology and mathematics but who are by no means experts in either field.

Once he has described the purpose of the chapter, Galton begins laying out his own theory of transmission of characters loosely based on Darwin's theory of pangenesis. According to Darwin's theory, the cells that are created in the zygote contain minute particles called "gemmules" each of which represents a different physiological feature inherited from the parent. As the cells multiply, they throw off these small particles which, "when supplied with proper nutriment, multiply by self-division, and are ultimately developed into units like those from which they were originally derived" (*The Works of Charles Darwin*, 321). Darwin suggests that the differential dispersion of gemmules could be responsible for variation and that this dispersion could be affected by damage to the reproductive organs caused by changes in their physical conditions and the effect of environmental conditions on the parents which would cause them to produce modified gemmules in adaptive response to these changes (Gillham 174).

Darwin's basic concept of the particulate gemmule is clearly echoed in the opening section of Galton's second chapter "Particulate Inheritance" in which he describes generally the process by which traits are inherited. With no concepts of genes—the term "gene" is not even coined until 1911⁷²—Galton makes the case that the fundamental unit of inheritance is particulate and sensitive to different environmental influences.

We seem to inherit bit by bit, this element from one progenitor that from another... while the several bits are themselves liable to some small change

⁷² "Gene," *Oxford English Dictionary* (Oxford: Oxford University Press, 2006) 27 Mar. 2006 <<http://dictionary.oed.com/>>.

during the process of transmission. Inheritance may therefore be described as largely if not wholly “particulate,” and as such it will be treated in these pages. (7)

In these lines which echo the essential tenets of Darwin’s theory of pangenesis, it is also possible to find similarities to the ideas of Mendel, whose work was not known by Darwin or Galton. Specifically, there is a comparable notion in Mendel’s work that organisms were made out of a mosaic of different traits and that these traits were particulate in that they could be represented as distinct entities.

Though Galton and Mendel share this common position on the separate particulate nature of traits, their concepts of individual traits are not the same. While Mendel saw the trait as a single cohesive fundamental unit, Galton argues that traits are made up of smaller particulate units. In the second chapter of *Natural Inheritance*, for example, he writes,

So in the process of transmission by inheritance, elements derived from the same ancestor are apt to appear in large groups, just as if they had clung together in the pre-embryonic stage, as perhaps they did. They form what is well expressed by the word “traits,” traits of feature and character—that is to say, continuous features and not isolated points. (8-9)

In some ways, the fact that Galton sees traits as a collection of characters rather than indivisible units does not necessarily make his position substantially different from Mendel’s. If, for example, similar elements cling together, then they represent in practice

a single cohesive unit. However, because they are divisible, this also means that there is the possibility of integration between them and the particles of the complementary trait from the other parent. This integrated or blended inheritance is similar to the kind described by Nägeli.

In chapter two of his book, Galton explicitly supports the possibility of both blended and mutually exclusive inheritance allowed by a multi-particulate character model.

There are probably no heritages that perfectly blend or that absolutely exclude one another, but all heritages have a tendency in one or the other direction, and the tendency is often a very strong one. By what we may see in plots of wild vegetation, where two varieties of a plant mix freely, and the general vegetation becomes a blend of the two, or where individuals of one variety congregate and take exclusive possession of one place, and those of another variety congregate together. (13)

In this metaphor of seed dispersal, Galton describes trait distribution in terms of separate particles which, like different seeds scattered on the wind, find purchase in different combinations on different plots of land. In some rare cases, however, the resulting flora is equally mixed, while in other cases one type tends to dominate over the others. As time wears on, however, organic stability may develop in those areas where some seeds tend to predominate. This creates a stable position for a particular trait.

Although stable positions can be accomplished, the degree to which these stable conglomerates affect reproductive outcomes is still influenced by the regular interference of chance. At the close of the chapter, Galton makes this argument explicitly. He also argues that the influence of chance in the hereditary process, based on an a priori probability, makes this process amenable to description by the mathematics of probability as they are instantiated in the law of errors.

What has been said is enough to give a clue to the chief motive of this chapter. Its intention has been to show the large part that is always played by chance in the course of hereditary transmission, and to establish the importance of an intelligent use of the laws of chance and of statistical methods that are based upon them, in expressing the conditions under which heredity acts. (*Natural Inheritance* 17)

Though Galton's vision of the process of trait distribution through reproduction includes random selection of traits and dominance, both of which are present in Mendel's theory, Galton's position on these two issues is not identical. First, Galton and Mendel's concepts of dominance are not identical. For Mendel dominance is an inherent feature of one character or another, whereas for Galton it is a position of stability among groups of particles that tends to gain a greater degree of permanence over time. Second, the role of random selection is different in each of their theories. For Mendel, the choice of which of the two traits from the male and the female make it into the genetic make up of any one particular offspring is random. Though Galton has a similar theory for the selection of particles, this theory differs to some degree because random selection can act in two

different ways. First, there is random blending of characters which can lead to a series of different possible character expressions. Second, there is the random selection of groups of particles which is similar to the phenomenon described by Mendel.

Galton's model of heredity, therefore, embraced both continuous and discontinuous variation. This model, which accepted all of the beliefs about hereditary outcomes that were accepted at the time he was writing, may have had a substantial influence on Galton's success with his critics, especially in making the case for the general applicability of his mathematical law of regression and his principle of geometrical diminution of characters. The opposite case can be observed with Mendel who was challenged by Nägeli on the grounds that he was presuming to make the case that discontinuous variation was the only form of inheritance despite the existence of facts which challenged this assertion.

Throughout chapter two, Galton offers a view of the general tendencies and operations of heredity limited to the process in which hereditary particles combine and the contributions of the parents are mixed. In this description, he makes the case for the application of the special topics of probability to the study of heredity based on the influence of chance on the selection and combination of characters.

Whereas chapter two restricts its scope of investigation to the mixing of traits in a single reproductive instance, chapter three offers the reader a broad multi-generational vision of reproductive outcomes. Despite the difference of its scope, Galton uses his examples, explanations, and arguments in chapter three for the same end as he used them in the second chapter: to make the case that probability plays a major role in the process of inheritance and, therefore, the mathematics of probability and statistics have

jurisdiction in making arguments about heredity. The change of scope, however, does impact the direction of Galton's argument because it allows him to argue for the necessity of taking a population approach, which Galton describes as a statistical approach, to the study of heredity.

In the opening paragraph of the chapter, Galton changes the scope of his investigation from the process of trait combination and selection to a description of the distribution of traits within a population.

The total heritage of each man must include a greater variety of material than was utilized in forming his personal structure. The existence in some latent form of an unused portion is proved by this power, already alluded to, of transmitting ancestral characters that he did not personally exhibit. Therefore the organized structure of each individual should be viewed as the fulfillment of only one out of an indefinite number of mutually exclusive possibilities. His structure is the coherent and more or less stable development of what is no more than an imperfect sample of a large variety of elements (*Natural Inheritance* 18).

In this opening explanation, a few elements are remarkable. First, Galton acknowledges that some traits are unexpressed which reveals that he, like Mendel, had something like an idea of the distinction between the genotype, the total number of characters that exist, and the phenotype, the number of characters which are physically manifested in the organism. Second, he makes the case that the elements which are expressed are only some out of an indefinite number of mutually exclusive possibilities.

This statement reveals that Galton was thinking along the lines of a gene pool of which individuals were only samples. This contrasts strongly with the traditional “essence” theories of species, such as the ones proposed by Kölreuter, Gärtner, and Quetelet, in which variation was only deviation from an ideal type.

Once he has made these general characterizations about the distribution of variation in populations, Galton makes specific arguments supporting his position. He begins by challenging the traditional model of inherited similarity as the direct transfer of traits by describing in detail what he believes is an equally plausible gene pool scenario.

It appears that there is no direct heredity relation between the personal parents and the personal child... but that the main line of hereditary connection unites the sets of elements out of which the personal parents had been evolved with the set out of which the personal child was evolved. (*Natural Inheritance* 19)

In this statement, Galton reveals that he is thinking of heredity in terms of populations rather than in terms of individuals because he sees parents and children as random expressions of characters out of a collection of traits. This characterization provides a rationale for accepting the necessity of statistics as a tool for studying inheritance. If heritability is seen as a fairly direct transfer of traits which are only slightly altered as a result of environmental or other factors, then the phenotype of the parents and the individual offspring can suffice as the subjects for investigating inheritance. If, however, individual offspring are simply one of infinitely many combinations of a set of heritable traits, then an investigation of the features of single individuals becomes

meaningless, particularly because there is no way of directly observing the combination of the particles themselves.

As a result, the study of inheritance by necessity becomes the study of fraternities or large samples of population which are capable of revealing patterns in and ranges of character combinations, a point also recognized by Mendel. This type of assessment, according to Galton, can only be accomplished through the use of statistical analysis.

We are unable to see particles and watch their groupings, and we know nothing directly about them, but we may gain some idea of the various possible results [of the combination of character traits] by noting the differences between the brothers in any large fraternity... whose total heritages must have been much alike, but whose personal structures are often very dissimilar. That is why it is so important in hereditary inquiry to deal with fraternities rather than with individuals, and with large fraternities rather than small ones. (19-20)

Once Galton has made this case for his “gene pool” model of inheritance and the subsequent need for a statistical approach the study of heredity, he presents his specific arguments about how stability of character is maintained in a system where traits are randomly selected from a large pool of possibilities. In this section, he turns to analogy to make his arguments. His choice to make arguments from analogy and hypothesized description rather than from the authority of previous experiments and observations suggests that Galton is trying to make the case for a relatively non-specialist audience

that his methodological approach is novel and more effective than previous approaches to the study of heredity.

He begins by positing that for all randomly combining particles there are tendencies of attraction and repulsion. He adds that over a period of time there will be a few groupings that will tend to reoccur and after multiple reoccurrences to become fixed as a stable pattern or form.

In every congregation of mutually reacting elements, some characteristic groupings are usually recognized that have become familiar through their frequent reoccurrence and persistence. Being less evanescent than other combinations, they may be regarded as temporary stable forms. (*Natural Inheritance* 22)

As evidence for his conclusion, Galton presents analogous instances in which constituent elements seem to tend towards certain forms. He draws on analogies from governments, crowds, and landscapes—all of which are composed of diverse elements that tend to commonly combine into certain stable forms such as democracy, an evening party, and “true to nature” artistic compositions. These analogies, although illustrative of the concept of stable form, seem quite unusual for describing a physiological phenomenon because they are social, emotional, psychological phenomena rather than physical ones. They seem to serve Galton’s purposes, however, because they are accessible to a wide readership and because they can be used to describe entities whose constituent parts can take multiple stable forms, which he calls “types,” and whose forms can change over time.

Throughout the second and third chapters where Galton makes the case for his theory of inheritance, there are only three citations of scholarly sources. There is one citation each in the field of anatomy and cytology and one from his own work on the faculties of twins. There are no sources from hybridization studies, animal breeding, or even from Darwin's theory of pangenesis, which Galton relies on heavily as a platform for his conceptual model, to provide evidence to support his claims. In each case where citations do appear, the sources cited do not supply evidence or ethical support for critical ideas in the text.

At the end of the third chapter Galton, who recognizes that this approach to developing his theory of heredity may appear dubious to the reader, provides a rationale for not drawing on the existent scholarship. He writes, "I have largely used metaphor and illustration to explain the facts, wishing to avoid entanglements with theory as far as possible, inasmuch as no complete theory of inheritance has yet been propounded that meets with general acceptance" (34).

By refusing to acknowledge the methods and results from previous approaches to the question of heredity, he is effectively casting them for the reader as largely inadequate and, therefore, obsolete. Having characterized previous scholarship as such, he can then position his own approach as a contrastively novel and effective alternative.

Rhetorically, the second and third chapters serve three important purposes in the overall argument presented in the text. First, they make the case for the collection of data on traits in large populations as the only adequate source of evidence for arguments about heredity. Second, they provide arguments supporting the jurisdiction of the special topics of probability, which represent the general mathematical foundations for the law of

errors, in making conclusions about the hereditary process and its results. Finally, they shape the reader's impressions about Galton's approach as novel and potentially successful for solving problems more traditional approaches have failed to solve.

Chapters Four and Five: Demonstrating Compatibility with the Law of Error

After describing his model for the distribution of traits and generally making his case for the application of probability and statistics to the study of heredity, Galton turns his attention to arguing specifically for the applicability of the law of error to investigations of the hereditary process. Chapters four and five involve particularly delicate arguments toward this end. Their delicacy results from the fact that Galton wants to make the case, on the one hand, that the principles and the formulae associated with the law of error are applicable to the study of heredity. On the other hand, he also wants to argue for a new way of interpreting the law of errors, which changes the focus of attention from the median value of a measurable trait to the deviation of trait values from the median, while at the same time maintaining that this interpretation does not violate the basic precepts of the law of error. In order to get his reader to accept his analogy between the law of error and the distribution of traits in a human population, Galton appeals both to his readers' subjective experience with trait variation as well as their own Victorian sensibilities about the importance of knowing one's place or rank in the scheme of things.

In chapter four, "Schemes of Distribution and Frequency," Galton begins his persuasive efforts by defining his new "science of heredity" and its practices, by laying

out for his reader the value of applying the law of error to statistical data, and providing a graphical demonstration of what the results of this application yield. Because he is introducing a new science, Galton begins by generally defining the objects of study for this science and the methods by which these objects are investigated. In the first sentence, he explains that these objects are, “fraternities and large populations rather than... individuals” (*Natural Inheritance* 35). In the second sentence, he lays out the methodology for the study of these subjects which he describes as, “a compendious method is therefore requisite by which we may express the distribution of each faculty among the members of any large group, whether it be a fraternity or an entire population” (35).

The “compendious method” that Galton alludes to but does not directly identify here is, of course, the law of errors. Though it may seem strange that he does not mention his methodological touchstone in the opening of the chapter, his silence here is an important strategic choice. The law of error, as it was applied by Quetelet and commonly used in physics, history, astronomy and sociology, was most often, if not exclusively, employed as a means of ascertaining the *mean* value of a particular feature in a population. Galton recognizes that this particular focus on the mean as the object of study is likely the one that his reader, if he has had any experience with the law of error at all, will recognize. In Galton’s unique application of the law to the subject of inheritance, however, it is the *deviation of a particular variation from the mean* not the mean itself that is the object of interest.

Because the primary focus and goal of the traditional use of the law of errors is to find the median value in a distribution of values, Galton begins his attack on the orthodox

application of the law by assaulting the notion that the median value provides the researcher with the most significant information about a population. To challenge the importance of the median and to persuade his reader that the range of variation is more important, Galton offers an example from the financial world. He describes a case in which the median income of the English population is known and explains that this information tells virtually nothing about the population itself. In contrast, a description of the deviation of a particular variation from the mean can reveal, “what proportion of our countrymen had just and only just enough means to ward off starvation, and what were the proportions of those who had incomes in each and every other degree” (Galton, *Natural Inheritance*, 35-36). In other words, knowledge of the mean and the standard deviation allows individual values to be ranked or compared with all other values. Although the mean alone can also serve as a ranking index, it can only supply an idea of whether a value is above or below normal, whereas knowledge of the standard deviation can reveal precise information about the percentage of the population which does or does not share a similar value.

The concept of rank within the overall population is important for Galton and most likely for his Victorian readers, who existed in a culture obsessed with the concepts of progress and betterment. Although Galton is a strict believer in the determinacy of natural endowment, rejecting much of the Victorian sentiment surrounding self-improvement, he makes a point at the end of his example to emphasize the importance of the concept of rank and its value in ascertaining social progress.

A knowledge of the distribution of any quality enables us to ascertain the rank that each man holds among his fellows, in respect to that quality. This is a valuable piece of knowledge in this struggling and competitive world, where success is to the foremost, and failure to the hindmost, irrespective of absolute efficiency.... When the distribution of any faculty has been ascertained, we can tell from the measurement, say of our child, how he ranks among other children in respect to that faculty, whether it be a physical gift, or one of health, or of intellect, or of morals. As the years go by, we may learn by the same means whether he is making his way towards the front, whether he just holds his place, or whether he is falling back towards the rear. Similarly as regards the position of our class, or our nation, among other classes and other nations. (*Natural Inheritance* 37)

In these lines, Galton explains that knowledge of the distribution of trait values could provide the reader with a way of ascertaining how his children, with regards to these highly esteemed attributes, would measure up to other children as well as supply him with a means of tracking over time the improvement or decline in these attributes. In the competitive late nineteenth-century environment, a method by which the rank of one's offspring could be ascertained and tracked would be of the greatest utility. By characterizing the usefulness of knowing the distribution of values in terms of the ranking of offspring, Galton both draws upon the values of the reader to support his case and establishes a link in his reader's mind between the law of error and the study of heredity which prepares the way for acceptance of his methodology.

Once Galton has preemptively made the case for the importance of determining the range of values for a given trait in a particular population over the importance of knowing the median value, he then introduces the law of error in the fifth chapter of the text titled “Normal Variability.” Here he presents arguments that the law of error can be applied to natural data but that a distinction had to be made between the traditional applications of the law of error and the approach he is employing.

In the opening of the chapter, Galton introduces the concept of deviation and makes the case that when natural data are marshaled their organization reveals regularity in terms of deviation from the median which is similar to the regularity observed in the deviations from the median in the mathematically derived law of error. In the first section, “Normal Variability,” Galton explains that the mathematically derived model for the probable distribution of errors reveals a scheme⁷³ of deviations which is fixed and which differs to some degree from the distribution of the values represented in his anthropometric data. By mathematically “smoothing” the curve, however, Galton claims he is able to make the deviations of the anthropometric data fall nicely into line with the scheme derived from the normal law of error. As evidence for this claim, Galton presents the values of deviation in the known tables of probability integrals, which describe the scheme of the law of error, side by side with the smoothed values for deviation in the anthropometric data for height, weight, arm span, etc. in Table 3 of the text and draws his readers’ attention to the visual evidence of their quantitative similarity,⁷⁴ “All the 18 schemes of deviation that can be derived from Table 2 have been treated on these

⁷³ For Galton the term “scheme” refers to, “a compendium of a mass of observations which, on being marshaled in an orderly manner, fall into a diagram whose contour is so regular, simple, and bold, as to admit of being described by a few numerals... from which it can at any time be drawn afresh” (49).

⁷⁴ See Appendix H

[smoothed] principles, and the results given in Table 3. Their general accordance with one another, and still more with the mean of all of them, is obvious” (*Natural Inheritance* 54). The close similarities between the values for deviation in the natural data and the distribution of deviation values calculated from probability provide sufficient evidence, according to Galton, that all of the mathematical formulae and principles that have been developed for describing the law of error can also be applied to describing the distribution of natural data.

Although by smoothing the values in the anthropometric data Galton has shown that the law of error is mathematically suitable for describing deviations, or errors, from a true value in anthropometric data, he has not made the case that that they are suitable for describing *variability* which is the focus of his hereditary investigation. As a result, he still needs to persuade his reader that these deviations can be interpreted in terms of natural variation. To accomplish this, Galton makes arguments from definition which attempt to establish a general similarity between the concepts of error and variation. According to the standard application of the law of error, the values of the observational results which fall above or below the mean “true value” of an observation represent a distribution of erroneous observations. By calculating the deviation of possible errors from the median, it is possible to estimate how far away in contrast with other measurements from the “true value” a particular measurement lies and, therefore, calculate the probability of its erroneousness. As a result, deviations describe, in the nomenclature of the law of errors, “probable errors.”

In the second section of the chapter titled “Normal Curve of Distribution,” Galton makes the case that because variations and errors are caused by similar phenomena, “the

combined influence of a multitude of accidents,” they can be reasonably classed together (*Natural Inheritance* 55).

But errors, differences, deviations, divergencies, dispersions, and individual variations,⁷⁵ all spring from the same kind of causes. Objects that bear the same name, or can be described by the same phrase, are thereby acknowledged to have common points of resemblance, and to rank as members of the same species, class, or whatever else we may please to call the group. (55)

Based on his position that “error” and “variation” are descriptors of the effects of the same cause, Galton argues that the mathematics which applies to the calculation of errors should also reasonably apply to the calculation of variability.

All persons conversant with statistics are aware that this supposition [that errors, differences, etc. are all the result of a multitude of small accidents] brings variability within the grasp of the laws of chance, with the result that the relative frequency of deviations of different amounts admits of being calculated, when those amounts are measured in terms of any self-contained unit of variability. (55)

Although Galton makes the case for identity between the terms “error” and “variation” for the purpose of convincing the reader that the law of error can be used to describe variability, he is quick to differentiate his use of the law of error to study

⁷⁵ Here Galton is using the figure congeries to persuade the reader that the terms “error” and “variation” share a similarity with one another by offering a series of possible subjects which connect the two terms.

variability from the traditional applications associated with it. To make this distinction, Galton relies on arguments from definition. He begins with the conclusion that because the science of heredity has different ends than those of mathematics, the specific manner in which terms are employed by the one is not particularly useful to the other.

It has already been said that mathematicians labored at the law of error for one set of purposes, and we are entering into the fruits of their labors for another. Hence there is no ground for surprise that their nomenclature is often cumbrous and out of place when applied to problems of heredity. This is especially the case with regard to their term of “probable error.” (*Natural Inheritance* 57)

Although initially Galton has no quibble with the law of error’s concept of “probable error” (**Q**), and describes it as co-referential with variability, at this point in the text, he assaults the term as misleading and in need of proper clarification. He argues that in common parlance most people would take the phrase “probable error” to mean the most probable error which, mathematically, would be the median error. The median error, however, is not an error at all, but instead the “true value” or the case of zero probable error. Because the term is so misleading, he claims that his recasting of the statistical phenomena of “probable error” as “probable deviation” is not only justified, but necessary. In this way, Galton is able to separate his terminology from the terminology commonly employed in the law of errors while at the same time maintaining the position that conceptually the law of errors is compatible with the study of variability.

Making the Case for Regression and the Geometric Diminution of Variation:

Blended Inheritance.

After spending chapters two through five arguing for the jurisdiction of statistics and probability in the study of heredity generally and the applicability of the law of errors to data gathered from the natural world specifically, Galton is ready to introduce his reader to the theoretical insights about heredity which he believes can be gained from the application of the law of error to anthropometric data on inheritance. These efforts begin in chapter seven, “Discussion of the Data of Stature,” where Galton describes the mathematical patterns that he has uncovered by applying the law of errors to anthropometric data on variation in height revealing *blended* inheritance. They continue in chapter eight “Discussion of the Data of Eye Color” where Galton makes the case using eye color data that the theoretical mathematical deductions he has established in the previous chapter are also applicable to the study of characteristics that are mutually *exclusive*.

In order to describe the laws of heredity, Galton lays out for his reader his work organizing and comparing data on the statures (heights) of co-fraternal offspring, with the statures of their mid-parents. He begins by defining the theoretical constructions “co-fraternal” and “mid-parents” and by making the case for the appropriateness of describing parents and offspring using these constructions. He defines co-fraternal as, “all the adult sons and transmuted daughters of a group of mid parents who have the same stature⁷⁶,” and the term mid-parents as “the average of two values; one the stature of the father, and

⁷⁶ The term “stature” used here by Galton is equivalent to the term “height.” I will use the two interchangeably throughout this section.

the other the transmuted stature of the mother” (88, 94). He explains that the arithmetical “transmutation” of the values for height in females into their male equivalents is appropriate because it makes the male and female values compatible so that either their contribution of traits or their acquisition of traits can be considered with statistical uniformity. He transforms the stature of mothers and daughters by “adding to each observed female stature... one inch for every foot” (6).

In addition to these theoretical constructions and transformations of the constituents of the observed population’s heights, Galton also describes constructions for the general population itself. Because he is comparing the values for median and deviation of heights in two groups, mid-parents and co-fraternities, he also has to establish a theoretical value for the median and deviation of heights in the overall “general” population which encompasses both groups. From these values, he can not only make the case that there is no one-to-one correspondence between the statures of the parents and the statures of the offspring, but also describe the ratio by which the offspring and parents vary from one another. He establishes the value of the mid-stature of the population (**P**) at 68.5 inches and the value of the variability within the population (**Q**) at 1.7 inches.

Once he has defined his terms and given the values for the median and variability for the whole population, Galton presents his results graphically to his readers and makes the case, using graphics as evidence and the method of regression analysis he developed earlier with his work with peas, that there is a general tendency for the stature of the offspring to regress towards the norm. The sample used to make his argument is a mix of data. One set of data comes from his Records of Family Faculty Data, which he calls his

R.F.F. data, gathered through a general offer of prizes for data on the stature of immediate family and relations. He also uses a second data set he calls “Special Data” which was collected by circulating cards to trusted correspondents asking them to provide the measure of stature among brothers. The R.F.F. data supply measurements on 928 adult offspring and 205 Mid-parents. The special data contained in aggregate the stature measures for 738 brothers (Galton, *Natural Inheritance*, 72-75 and 78-79). Although the measurements from his anthropometric laboratory also included height, they did not specifically include the height of the brothers and parents of an individual. As a result, these measurements were not used in his arguments.

Galton begins his discussion of his regression calculations of the R.F.F. and special data by stating the conclusion he draws from it: “However paradoxical it may appear at first sight, it is theoretically a necessary fact, and one that is clearly confirmed by observation, that the stature of the adult offspring on the whole must be more *mediocre* than the stature of the parents” (95).

After defining his position, Galton provides his reader with a graphical presentation (Figure 1) of the data.

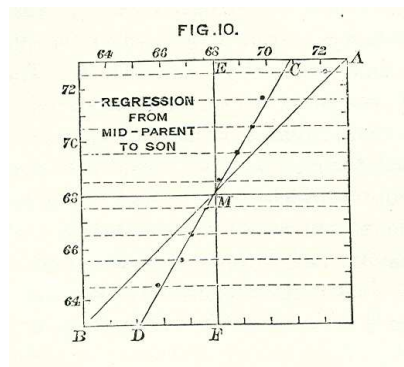


Fig. 1. Regression from Mid-Parent to Son⁷⁷

⁷⁷ *Natural Inheritance* p. 96

In this graphic, the stature (heights) of the mid-parents is listed along the left side and the heights of co-fraternities are listed along the top. Dots are placed at the point where the statures of the mid-parents, plotted vertically, and the average statures of the co-fraternities, plotted horizontally, intersect. Galton then draws line **CD** through these dots which represents the average relationship between the values of the mid-parents and their co-fraternal offspring. Next, he draws a second line **AB** which describes, “what the mid-statures of the sons would be if they were on average identical with those of their mid-parents” (*Natural Inheritance* 97).

After establishing that the statures of offspring are different from the statures of their parents, Galton makes the case geometrically that the difference between the two has a fixed value. By drawing a vertical line **EMF** through the place where **AB** and **CD** intersect, it is possible to discover the degree of difference between the stature of the offspring and the stature of the mid-parent by taking the ratio of **EA** to **EC** which is constant throughout the areas **EMA** and **EMB**. Galton reckons that the ratio between **EA** and **EC** is 2:3 which he calls the ratio of “filial regression” or, “the proportion in which the son [co-fraternal group] is, on the average, less exceptional than his mid-parent” (97). Because Galton believes that parents contribute equally to the stature of their offspring, “the average regression from the parental to the mid-filial [co-fraternal] stature must be the one half of two-thirds, or one-third” (98).

After making the case using regression that the median values of stature for groups of co-fraternities regress towards the mean by $2/3$, Galton makes the same case geometrically from the arrangement of the data on his results table. He explains that

while looking at the values as they were listed in rows and columns in table 11, it occurred to him that there seemed to be a relationship between values on the chart and those values which were near to them. After he smoothed the entries, “by writing at each intersection between a horizontal line and a vertical one the sum of the entries in the four adjacent squares,” he noticed that, “the lines drawn through the entries of the same value formed a series of concentric ellipses” (100-01). The places where, “each successive ellipse was touched by a horizontal tangent, lay in a straight line that was inclined to the vertical in a ratio of 2/3, and those where the ellipses were touched by a vertical tangent, lay in a straight line inclined to the horizontal in the ratio of 1/3” (101). In the midst of the proceeding description, Galton provides his reader with the following graphic (Figure 2).

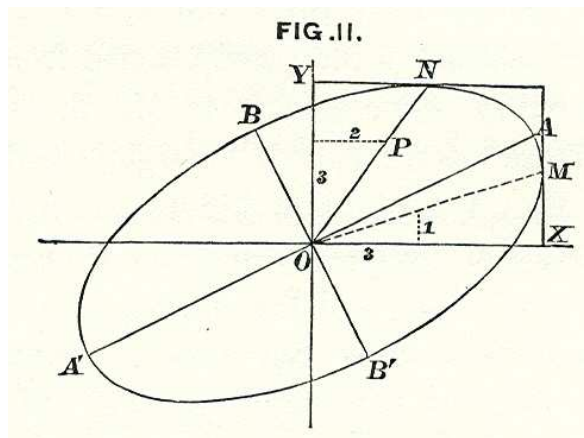


Fig. 2. Geometrical Proof of the Filial Regression Value from Natural Inheritance⁷⁸

⁷⁸ *Natural Inheritance* p. 101

This visual represents the resulting transformation of organized raw data into geometric proportions. By tracing these geometric proportions from the contours of the raw data as it was organized, Galton is again able to bring geometric principles to bear on the data in order to provide proof for his conclusions. In the text, however, Galton does not do a sufficiently lucid job of explaining the origin of the graphic nor of reproducing it in such a way as to clearly reveal the origin of its important elements. In an earlier article “Regression Towards Mediocrity in Hereditary Stature” in *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, Galton is much clearer both graphically and in writing about how this particular geometric figure emerges from his data.

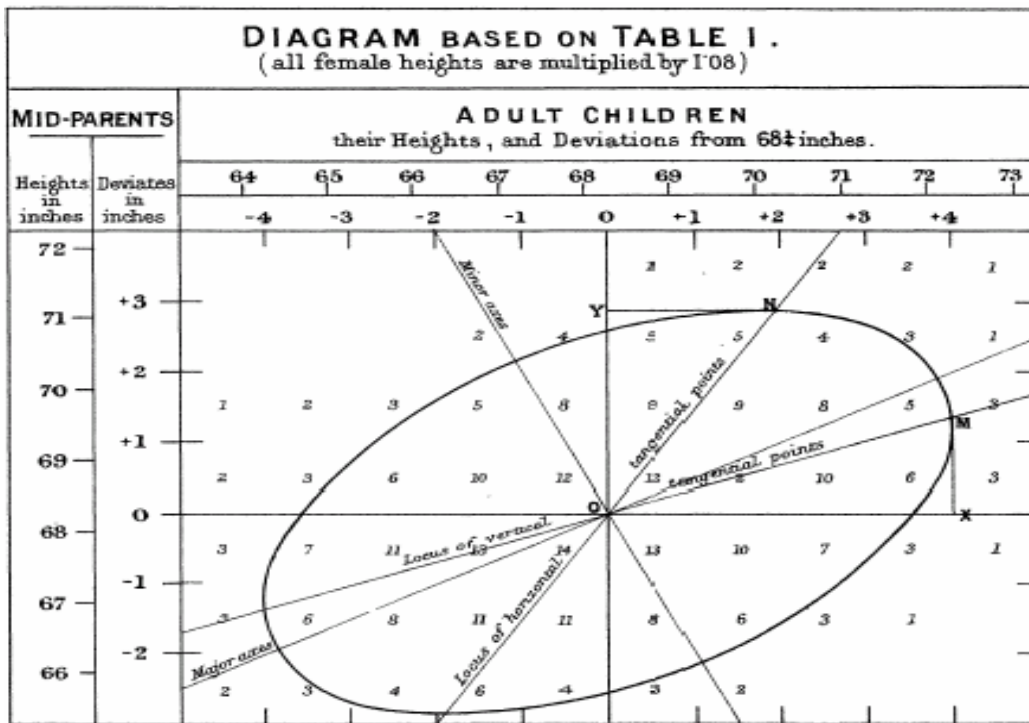


Fig. 3. Geometrical Proof of the Filial Regression Value with the Actual Data⁷⁹

⁷⁹ Francis Galton, “Regression Towards Mediocrity in Hereditary Stature,” *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 15 (1886): Plate X 248-49.

In the two sections of Galton's argument described above, he relies on proofs based on arithmetic and the axioms of geometry to make the case for the value of regression. Because these proofs could be deduced using geometry and arithmetic from both the graphical representation of his data and from an imposition of a geometrical arrangement on the physical data, his evidence for regression in this section is presented with multiple sources of credible deductive support.

Despite the strength of this evidence, however, Galton's arguments might be undone if the evidence cannot be shown to distribute according to the law of error and if it cannot stand up to the scrutiny of professional mathematicians. In order to avoid both of these problems, Galton has his results checked by sending his values for the variability of the general population and the co-fraternity, along with the average value of regression from mid-parent to co-fraternity to Mr. J.D. Hamilton Dickson, Tutor of St. Peter's College, Cambridge, to calculate their distributions according to the law of error. The results which he receives, though not exactly the same, are similar enough to his own for him to proclaim that they correlate with the law of error.

It is obvious from this close accord of calculation with observation, that the law of errors holds throughout with sufficient precision to be of real service, and that the various results of my statistics are not casual and disconnected determinations, but strictly interdependent. (*Natural Inheritance* 103)

On the basis of this success, he describes his findings from this point on as the “law of regression” which holds that, “the deviation of the sons from **P** [the mid-stature of the population] are, on average, equal to one-third of the deviation of the parent from **P** [The population mid-stature], and in the same direction” (104). In other words, the offspring have only $1/3$ the amount of variability in any given character of the combined average of their parents for the same character, their character values always being closer to the population mean. Following this assertion, Galton is careful to explain that his assumption of this law is justified because he can prove it using multiple deductive methods (an approach that Mendel takes in making his arguments for the law of development), because this law has been shown to hold in humans as well as peas, and because it holds for both blended and exclusive heredity.

Once Galton has established his law of regression, he then uses it as a jumping off point for making the case that variation in character traits diminishes geometrically between each generation. He begins by reasserting his finding that offspring regress from their mid-parents towards the norm by $2/3$. From this fact he reasons the amount of influence that the variation of one generation has on another is $1/3$.

Once he has posited this value of $1/3$ for inherited peculiarity (the value of variation from the population mean), Galton inquires as to what the total inheritance for a particular trait in a given generation would be if the value of $1/3$ remained constant in every generation. In order to investigate this, he imagines what the value of transmittable deviation would be for a given mid-parent if the value of deviation for each preceding generation was $1/3$. By adding these values to the degree of variation existent in the mid-

parent, which is one because it represents an undivided store of variation, he is able to calculate the amount of the Mid-parents bequeathable variation: $3/2$.

Call the peculiarity [value of variation from the population mean] of the Mid-Parent \mathbf{D} , then the implied peculiarity [value of variation from the population mean] of the Mid-parent of the mid-parent, that is the mid-grand-parent of the man, would on the above supposition be $1/3\mathbf{D}$, that of the mid-great-grand-parent would be $1/9\mathbf{D}$, and so on. Hence the total bequeathable property would amount to $\mathbf{D}(1+1/3+1/9+\&c.) = \mathbf{D}3/2$. (*Natural Inheritance* 134)

Even though it is a mathematically viable expectation that if each offspring receive $1/3$ of the variation of a particular character trait from their ancestors over time this variation could add up to a value of $3/2$, Galton argues that the law of regression strictly limits the amount of variation that each generation can pass down to $2/3 \mathbf{D}$ (the value of the mid-parent's variation).⁸⁰ Galton further offers a series of suppositions that he admits have no empirical support but which he believes must hold true given the law of regression and the generally accepted principle of reversion.

His first supposition is that the variation between generations should always be diminished by the same amount. Given this parameter, he argues, we should expect that if the value of the bequeathable variation from the mid-parent to the child without diminution is $3/2\mathbf{D}$ and the law of regression proves that the value of a particular

⁸⁰ The total amount of heritable variation in the mid-parent must be $2/3$ because it is cut in half when passing to the next generation making the actual amount inherited $1/3$ as required by the Law of Regression.

characteristics variation is actually $\frac{2}{3}D$, then the value of variation must have been diminished by $\frac{4}{9}$ or approximately $\frac{1}{2}$ in each generation.

His second supposition asks the reader to accept that for each generation further back variation shrinks geometrically. As a result, “the influence... of the mid-parent may be taken as $\frac{1}{2}$, and that of the mid-grand-parent by $\frac{1}{4}$, and so on” (136). Interestingly, this supposition, despite the lack of evidence, proves to be correct on the strength of the Mendelian model of trait distribution in which half the total endowment of genes of a parent is transferred, in turn transferring $\frac{1}{4}$ of the endowment of the genes transferred to the parent by their parents, and so on. In addition, the argument echoes Darwin’s argument for competition based on Malthus’ argument of geometrical versus arithmetical progression.

Galton’s final supposition is that it is possible that some remote ancestor could contribute more of a particular character than their mid-parent. This supposition accounts for the widely accepted phenomenon of reversion, whereby an offspring resembles more closely a grandparent than a parent (Bulmer 110). Though Galton recognizes the phenomenon of reversion, he argues that it should be considered an exception from the general process characterized in the first two suppositions.

In this final section, Galton argues for a geometrical diminution in traits over time. He makes his case from the law of regression and several hypothetical suppositions which have no empirical basis. Because of the thinness of his case, Galton does not explicitly argue in this text that this geometric progression represents a hereditary “law.” It is not until his 1897 article in *Nature*, “The Average Contribution of Each Several Ancestor to the Total Heritage of the Offspring,” nine years later that Galton is finally

able to dub this mathematical pattern of trait contribution the “law of ancestral heredity” after finding the empirical evidence he needed in breeding data describing the coat color of basset hounds.

In the seventh chapter of the text, Galton finally makes his case that by applying the concept of standard deviation from the law of error to anthropometric data and by comparing the standard deviations of two generations to one another using regression analysis, mathematical expressions for the patterns of probable similarity and difference between generations can be established. First, he uses his stature data to establish a mathematically describable law of regression using the special topics of geometry and the law of errors. Then, based on this law, he makes the case for the geometrical diminution of traits from one generation to the next.

Predicting Regression: Exclusive Inheritance

Although Galton makes a strong case for his patterns of heredity using empirical data and deductive arguments, he understood, as Mendel did, that in order for his mathematically describable laws to be accepted as truly rational and deductive they had to be able to predict the distribution of hereditary characters. In addition, he also realized that for his regression law to be considered a hereditary law, it should apply to both of the recognized types of inheritance, blended and exclusive. In this way, his work was more comprehensive than Mendel’s which did not seriously address the cases of blended inheritance which ran contrary to his position. In order to make the case that his laws were all inclusive and not only empirical but rational, Galton presents his reader in

chapter eight, “Discussion of the Data of Eye Color,” with a set of values for the distribution of eye colors predicted from his mathematically based laws and compares them to the actual distribution of eye color as described in his anthropometric data.

Galton begins the chapter by discussing the source of the data and the methods he used to gather and organize it. According to Galton, the eye color data also came from the “Records of Family Faculties” (R.F.F.) of 928 adult offspring and 205 Mid-parents. Results from the “Special Data” and the anthropometric laboratory were not used because neither of these data sets contained information on eye color.

Unlike stature, eye color is qualitative and, therefore, required Galton to construct a special scheme to define the different categories of eye color. Given the data he identifies eight distinct categories 1) light blue; 2) blue, dark blue; 3) grey, blue-green; 4) dark grey, hazel; 5) light brown; 6) brown; 7) dark brown; and 8) black (*Natural Inheritance* 141).

In addition to dividing up his data by eye color, Galton also separates his information by generation. To accomplish this, he creates four separate categories of heredity: category I represents the children of a pedigree (clearly identified extended family); category II the children’s parents, uncles, and aunts of the pedigree; category III the children’s grandparents, great uncles, and great aunts; and category IV great grandparents etc. (140). He uses these categories, along with the types of eye color, to organize the data in Table 15 of the text.⁸¹ In this table, Galton provides readers with the number of subjects in each generation with a particular eye color as well as the percentage of each eye color in a generation. His results reveal that the total frequency of any given eye color in proportion to the frequency of any other eye color does not

⁸¹ See Appendix I

significantly change from generation to generation. This reveals that there is segregation rather than a blending of traits from one generation to the next. Thus, eye color is exclusive inheritance in contrast to height which is blended.

Having established for each generation the percentage of the population which is endowed with a particular eye color type, Galton calculates, using his principle of the geometric diminution of character, the theoretical distribution of eye color in a particular generation. The results are listed by Galton in Tables 19⁸² and 20⁸³ in which he presents both the observed and the calculated eye colors for groups of families (Table 19) and for individual families (Table 20). The values in these tables are arrived at by calculating the eye color of the offspring based on the eye colors of the parents (I), the grand parents (II), and a combination of the parents and grandparents (III).

He begins by predicting the eye color of the offspring based on the eye color of the parents. Employing the law of regression and the tendency for the geometric diminution of traits over generations, he calculates that each parent separately must contribute 30% of the child's heritage: this calculation is based on the assumption that the average contribution of variation from both parents according to the law of regression is 1/3. If one parent contributes 30% of the child's heritage, then combined they should make up 60%. That leaves 40% to be contributed from other ancestry according to the tendency for the geometric diminution of traits (Galton, *Natural Inheritance*, 149).

Galton presents these contributions in percentages rather than fractions because unlike blended traits such as stature, in which a certain fraction of the final trait's makeup can be contributed by the parent, exclusive traits are either all or nothing. As a result, he

⁸² See Appendix J

⁸³ See Appendix K

needs to speak of them then in terms of the probability with which they might be expressed in a particular offspring or set of offspring. According to his calculations, there is a 30% chance of getting a particular feature from one parent or the other, and a 40% chance of inheriting a feature from a distant relative about whom nothing is known except that this relative belongs to the same race as the subject. In addition, these percentages also describe the percentage of the total population of offspring for a particular set of parents that should have a particular feature: for example, 30% of all offspring should have the features of one parent, 30% should have features of the other.

Once the general percentages for inheritance are assigned to a particular generation of offspring, these percentages must be, “rateably assigned between light and dark eye color, in the proportion in which those eye colors are found in the race in general,” or in the particular generational category in specific (150). If, for example, the 40% probability of inheriting a particular eye color from some distant relative belonging to the same race was divided according to the different percentages for light and dark eyes in the total race, 61.2% and 26.1% respectively, then the chance of getting light or dark eyes from this possible source for inheritance is, according to Galton, 28% for light and 12% for dark. It should be noted that in these calculations the general percentage of light and dark eye colors (87.3%) does not account for the total amount of inheritance from the race whereas the percentage of light and dark eye colors (40%), which represents the possibility of inheriting a particular trait from the general race, does. This discrepancy is the result of Galton looking upon hazel eyes as a blended inheritance of $\frac{2}{3}$ light and $\frac{1}{3}$ dark eye color. In order to simplify his inquiry into alternating inheritance, Galton decided to allocate the percent value for hazel eyes, which he lists on

the chart as 12.7% of the total race, in a 2:1 ratio so that out of the total contribution of 30% that hazel makes of the 40% of possible inheritance, 20% is allocated to the total for light eyes and 10% to dark. In other words, the existing case of blended inheritance, hazel eyes, in the predominantly non-blended hereditary category of eye color is made to support the 2:1 ratio predicted for its non-blending counterparts, light and dark eyes. This resolves the complication of the existence of blended inheritance in a population which Galton is trying to use to establish the verity of the law of regression for exclusive inheritance.

The calculations previously described, as well as those for the total percent of either light or dark eye colors for grandparents and for the combination of parents and grandparents, are all organized by Galton in Tables 17 and 18.⁸⁴ Following the theoretical calculation of these values using the law regression and the geometric tendency for trait diminution, Galton compares his calculated values to the actual data gathered on the eye color for different generations in Tables 19 and 20. The close similarity between the values calculated from the laws and the values observed in nature encourages Galton to exclaim, “A mere glance at Tables 19 and 20 will show how surprisingly accurate the predictions are, and therefore how true the basis of the calculations must be” (*Natural Inheritance* 152).

Despite the promising outcome of his results, however, Galton tempers his claim with the recognition that his sample is statistically too small to provide robust enough results for him to claim complete success in his investigation: “My returns are insufficiently numerous and too subject to uncertainty of observation, to make it worth while to submit them to a more rigorous analysis” (153). Nevertheless, he considers them

⁸⁴ See Appendix L

robust enough to support his position: “but the broad conclusion to which the present results irresistibly lead, is that the same peculiar hereditary relation that was shown to subsist between man and each of his ancestors in respect to the quality of stature, also subsists in respect to that of eye color” (153).

With these words, Galton concludes the final portion of his text in which he lays out for his readers evidence that his laws of heredity and regression have the power to describe if not predict the process of heredity. Though there are three more substantive chapters, on artistic faculty, disease, and latent elements (i.e. the characteristics not expressed in the genotype), and a summary which follows, they do not substantively build the theory beyond the point to which he has taken it in the eighth chapter. But they do show a willingness to extend the discussion of heredity to behaviors and propensities and not just observable physical traits.

Combined, chapters one through eight represent a deliberate and well orchestrated argument in which Galton makes the case for a mathematical approach to the study of heredity. Within these chapters, Galton argues generally for the jurisdiction of probability and statistics in the study of heredity, advocates specifically for the applicability of the law of errors to anthropometric data, draws conclusions using the special topics of arithmetic, geometry, and the law of errors that patterns exist in the hereditary process, and provides evidence presented as compelling that these mathematical patterns are supported by empirical evidence.

The Reception of *Natural Inheritance*

The reactions of reviewers in professional journals and in widely-read public newspapers and magazines to Galton's statistical-probabilistic approach to the problems of heredity provide evidence that the tactics he used in *Natural Inheritance* were effective in persuading his readers 1) that he was taking a novel approach to the study of heredity, 2) that statistics and probability had jurisdiction in and were effective methods for investigating heredity, and 3) that regression and geometrical diminution were fundamental principles of heredity. In addition, they reveal the existence of a readership which both understood the principles of probability and statistics and which was favorably disposed to their application in biology.

That Galton's arguments for the novelty of his work in chapters two and three of the text were successful is clear in reviewers' comments, particularly in their assertions that Galton's work dealt with some important issues about heredity raised by the work of Darwin that had not been sufficiently investigated by his predecessors. This position is expressed by the mathematician John Venn, originator of the Venn diagram, in his review of the book in the journal *Mind*.

This seems to me a very important contribution to a neglected side of the doctrine of heredity or descent. It is an attempt to apply accurate quantitative methods to the various successive steps by which one generation of organized beings follows another. Statistics in abundance have long been available as to the characteristics of each such generation separately.... But no one so far as I know, had hitherto

thought of tracking the intermediate steps, and of raising and answering the question... Why it is that successive generations thus continue to resemble each other (Venn 414).

Another response suggesting that Galton had pushed the boundaries of inquiry about heredity in an important direction appeared in the *Scottish Leader* in which Professor Patrick Geddes observed that Galton “had the subject of heredity all to himself” while “most eminent biologists, from Huxley downwards, despite their strong Darwinian faith and advocacy have... stuck to pre-Darwinian problems instead of attempting the solution of the far more important post-Darwinian ones” (Gillham 267).

In addition to being convinced that Galton was exploring new topical territory in his work, there was also a sense among his reviewers that they were persuaded that his work offered a new and relevant *methodological* approach to the subject as well. This opinion is expressed in Venn’s comments and by an anonymous reviewer in *Science* who writes, “A prominent feature in this work is the application of ‘the probability curve’ to the facts of physical variation—an attempt to apply mathematical conceptions in the field of biology, and to found a science of biological statistics” (“Natural Inheritance” 322). In this review, the anonymous reviewer not only expresses his agreement with Galton’s contention that his use of the principles of statistics and probability to investigate biology are novel, but he also accepts Galton’s argument that the special topics of probability and statistics have jurisdiction in biological investigations. These convictions are shared by John Dewey in his review of *Natural Inheritance* in *Publications of the American Statistical Association*.

This work is of double interest. Its primary purpose is biological, being to subject the question of heredity to accurate quantitative and mechanical treatment. As such it is doubtless the ablest work on the subject extant. But in the course of his investigation Galton has collected a large mass of statistical information, and, what is more important, has developed some new and interesting statistical methods. (Dewey 331)

In addition to providing evidence for the overall success of Galton's persuasive efforts in the text, the reviews also offer proof that there were groups in his readership who understood and appreciated his statistical/probabilistic approach to heredity. All of the reviewers in the scientific journals, for example, discussed the mathematics in detail which suggests that at least in scientific professional circles there were some scientists who had been exposed to and understood statistical theory (Dewey 1889, "Natural Inheritance" 1889, Venn 1889). That Dewey, who was a mathematician writing for a statistical journal, understood the mathematical intricacies of the argument is no surprise at all.

That the scientific reviewers in *Nature* and *Mind* could write in detail about the mathematics is not as significant, perhaps, as the fact that they chose to write at length about the mathematical aspects of *Natural Inheritance*. By extensively discussing these aspects of the argument, the reviewers indicate that they thought their respective audiences in the sciences were ready and, perhaps even able, to appreciate the mathematical intricacies of Galton's argument.

Although the scientific and mathematical reviewers discuss Galton's mathematical arguments in depth, the reviews in the popular circulars, such as the *Scottish Leader* and the *Spectator*, only briefly mention the mathematical elements. The reasons for the scant treatment are most likely 1) that the reading publics that they were addressing were not familiar with or not generally interested in the mathematics of probability and statistics; 2) that unlike the scientific journals they did not have as much space to write long explanatory pieces; and 3) that the reviewers in popular publications were less familiar with the mathematics of statistics and probability than the scientists and mathematicians and, therefore, less comfortable writing in detail about them.

A common theme running through the commentaries in the popular circulars was that Galton's text presented a novel and important theory which was, however, difficult to access even for well educated members of the public. The critic in the *Spectator* writes that the text was difficult, "even for those competent to understand it," and that it required, "careful and concentrated study," in order to comprehend it. Despite these difficulties, however, he writes, "it is well worth the time and trouble needed to master it. It lays the foundations of what one day will be a great science" (Gillham 267).

An analysis of the reviews of *Natural Inheritance* reveals that an audience which understood and was receptive to the arguments from probability and statistics existed and that for this audience the content and method of Galton's arguments were novel and appropriate for making his case about the process of heredity. In the decade which followed the publication of *Natural Inheritance*, the feelings expressed by the critics seemed to be shared by biologists as well. Studies of the process of variation such as William Bateson's "Some Cases of Variation in Secondary Sexual Characters

Statistically Examined” (1892)⁸⁵ and W.F.R. Weldon’s “On Certain Correlated Variations in *Carcinus moenas*” (1893)⁸⁶ began to emerge which employed Galton’s biometrical method of gathering and graphically organizing data for the purpose of studying the change in the frequency of characters in populations. The fact that Galton’s work was not only praised, but also emulated suggests that the time had arrived for theories which applied the special topics of probability and statistics to the study of heredity.

Conclusion

Whereas Darwin used quantification and quantitative comparison as a source for support and discovery in his arguments and Mendel appealed to the mathematical principles and operations of probability and combinatorics to design his experiments and support the lawfulness of his conclusions, Galton was the first to use an existing set of mathematical principles and operations as an organizing analogy for understanding natural phenomena. Although it is impossible to know what specific elements or combination of elements to credit with the success of Galton’s argument, this investigation of the rhetor, the argument, and the audience reveals several factors which most likely played a central role in its triumph.

One factor influencing the success of Galton’s arguments was his position within the scientific elite of England. Although Mendel’s theory of heredity was also novel in

⁸⁵ William Bateson, “Some Cases of Variation in Secondary Sexual Characters Statistically Examined,” *Proceedings of the Royal Society* 54 (1893): 585-94.

⁸⁶ W.F.R. Weldon, “On Certain Correlated Variations in *Carcinus moenas*,” *Proceedings of the Royal Society* 54 (1893): 318-33.

important ways, Galton had a much better chance of being taken seriously because he had established his scholarly credibility through his earlier work and because he was a well-known figure with social connections within different scientific institutions. This contrasts starkly with the social and academic situation of Mendel who had a sparse and unremarkable record of scholarship before he published his “Experiments” and who was a virtually unknown figure living far away from the social and academic centers of science in Western Europe.

A second factor contributing to the success of Galton’s argument was that unlike Mendel his basic conception of hereditary was not in conflict with the established wisdom on the subject. Whereas Mendel’s arguments and evidence embraced the position of non-blended discontinuous inheritance exclusively, Galton’s work allowed for the possibility of both blended, continuous inheritance and exclusive, discontinuous inheritance. That Galton’s inclusiveness played well with his readers is evident in the lack of criticism about the substance of his position and the success he had in persuading his readers that his results were generally applicable across all cases of heredity in many species. Mendel, on the other hand, is ignored by all except Nägeli, whose criticisms were aimed at Mendel’s insistence on the existence of only discontinuous inheritance and focused on chastising him for assuming that what he had found to be true in some traits in peas was true in all traits in all organisms.

A third factor influencing the success of Galton’s work was the success of his initial persuasive arguments deployed in his first chapters to justify his application of the methods and principles of statistics and probability to the study of heredity. The importance of this factor is supported in the reviewers’ acceptance and even approbation

of Galton's use of probability and statistics to develop biological laws. Galton's success in this persuasive endeavor probably had something to do with his use of readers' own subject experience with human characteristics as well as their cultural predisposition to be concerned with rank and order as common ground for getting them to accept the law of error as an organizing analogy for understanding the process of heredity. In Mendel's work, there was no overt attempt to persuade his audience that statistics, probability, and combinatorics could be used to describe the distribution of traits and the character of the reproductive process: even if there had been, it would have been difficult to construct as persuasive an analogical argument using peas.

With good established ethos, a popular position on inheritance, and a subject of study that his audience could identify with, it may, perhaps, come as no surprise that where Mendel's brilliant arguments failed, Galton's attempts to persuade readers to accept the use of the special topics of probability and statistics to investigate the invisible processes of heredity succeeded. In the end, Galton's success proved to be a crucial factor in the revival of Mendelian inheritance because in the decade following the publication of *Natural Inheritance*, Galton's statistical probabilistic methodology for the study of heredity rose to prominence through the work of William Bateson, Karl Pearson, and Frank Weldon—all of whom would develop and provide further evidence to support the efficacy of his approach. It is within the context of this growing mathematization of biology and the debates which arise between biologists about the continuous or discontinuous nature of inheritance that Mendel's monograph resurfaces and ushers in the modern study of genetics.

Chapter 5: Behind the Curve: The Place of Mathematics in Making Claims about Heredity at the Opening of the Twentieth Century

But I am horribly afraid of pure mathematicians with no experimental training. Consider Pearson. He speaks of the curve of frontal breadths, tabulated in the report, as being a disgraceful approximation to a normal curve.... This skew curve fits the dozen observations at the two ends better than a normal curve; it fits the rest of the curve, including 90% of the observations, worse. Now this sort of thing is always being done by Pearson, and by any 'pure' mathematicians. —Frank Weldon to Francis Galton⁸⁷

The development of mathematical methods by Professor Karl Pearson and a few others promised for a time the key to the riddles of evolution. This movement, led by a pure mathematician, developed a series of beautiful methods for the mathematical analysis of data and the comparison of variations. These methods are of the greatest importance when rightly used. —George Shull on Karl Pearson⁸⁸

With his arguments in *Natural Inheritance*, Galton initiated the development of a mathematical theory of inheritance and variation that drew on the principles of the error curve and employed statistical methodology for gathering and describing biological phenomena in populations. This approach to heredity was adopted by a small but important group of young natural researchers in the last decade of the nineteenth century who began to use the error curve to invent and support arguments about the nature of variation, evolution, and heredity.

During this period, three of Galton's most notable disciples—Karl Pearson (1857 – 1936), William Bateson (1861-1926), and Frank Weldon (1860-1906)—increased the number of biological phenomena investigated using Galton's techniques and developed novel mathematical applications for biological research. While Bateson and Weldon represented a new breed of biologist who adopted the tools of statistics and probability in

⁸⁷ Robert Olby, "The Dimensions of Scientific Controversy: The Biometric-Mendelian Debate," *British Journal for the History of Science* 22 (1988): 305.

⁸⁸ Geo H. Shull, "Heredity as an Exact Science," *Botanical Gazette* 50.3 (1910): 226.

their efforts to understand biological phenomena, Karl Pearson was an even rarer creature: a mathematician interested in biology for the sake of expanding mathematical applications. Inspired by the efforts of Weldon and Galton to apply statistics and probability to biological phenomena, Pearson developed not only new statistical tools for graphically describing variation and evolution but also a novel biometric theory of heredity based primarily on the principles of the error curve and correlation.

This chapter investigates two important rhetorical issues: 1) to what extent mathematics was accepted by biologists studying variation, evolution, and heredity in the last decade of the nineteenth century, as a legitimate source for arguments about these organic phenomena? And 2) to what degree was the acceptability of a mathematical line of argument contingent on the acceptability of a pre-existing biological line of argumentation? In this chapter, I will make the case that during the last decade of the nineteenth century there was a general feeling in the biological community that the biological arguments from the special topics of probability and statistics developed by Galton in the previous decade were legitimate sources for arguments about variation, evolution, and heredity. In addition, I will argue that though mathematically informed arguments were generally accepted, the acceptance of a particular line of argument from mathematics depended on the audience's willingness to accept the biological model upon which the mathematical arguments were based.

Karl Pearson (1875-92)

While Francis Galton is considered the conceptual father of biometry, Karl Pearson might be best described as its first radical disciple. With his zealous, some would say single-minded, efforts and considerable mathematical talents, he became a major force, along with Frank Weldon, in developing and promoting the use of mathematical models to describe the action of evolution in populations of organisms.

In recent years Pearson and his work have been getting more attention from historical scholars because of their importance in initiating the development of a mathematically rigorous program of statistical biology in the early twentieth century. In the first extensive biography of Karl Pearson, Theodore Porter suggests that Pearson's drive to expand the purview of mathematics to include biological phenomena was influenced by his belief in the humanist ideal of the broad pursuit of knowledge and his opinion that a broad base of theoretical knowledge could serve as a source of inspiration for finding solutions to practical problems. His belief in humanist education and the importance of theoretical knowledge is evident in the range of his interests in his formative college years, in his later work as a teacher, and in his contributions to the physical and biological sciences.

In his years as a student and youthful traveler (1875-80), Pearson's activities indicate his multi-faceted interest in language, culture, politics, and literature as well as a focused attention on the study of mathematics. From early on, Pearson was encouraged to study mathematics by his father William Pearson.⁸⁹ In 1875, before he was even admitted to college, his father hired a famous tutor, E.J. Routh, to prepare Karl for competition in the prestigious Cambridge mathematical Tripos, a grueling nine-day mathematical

⁸⁹ A prominent lawyer who garnered the title of "Queen's Council" in 1875. (Porter 16)

examination whose winner garnered perhaps the highest academic prestige that could be bestowed on an undergraduate at Cambridge (Porter 22, 40).

Though Pearson's father focused Karl's ambitions on mathematics, Karl directed his own energy into the study of the humanities, especially literature. In 1877 he began a commonplace book in which he included commentaries on many of the authors he was reading including Shelley, Rousseau, Voltaire, and his favorite, Goethe (27). Porter argues that these choices were most likely influenced by the friendship that had developed between the young student and the celebrated English literary figure, Oscar Browning.

Pearson's pleasure reading and his relationship with Browning, according to Porter, moved his interests toward German philosophy, most importantly the phenomenological philosophy of Ernst Mach. In the summer of 1877, Pearson traveled with Browning to France and Germany. In the year following his return, Pearson read heavily in German philosophy and began sharing his opinions openly on German philosophy and literature with fellow students Robert Parker, John Lawrence Green, and Edwin Cooper Perry. Based on his regular conversations with his fellow students, Pearson began to develop a "Teutonic metaphysical idealism" which, Porter explains, joined a search for individual meaning with a growing social concern for improving the lot of the laboring poor (Porter 33).

Despite his blossoming interest in literature, philosophy, and all things German, Pearson continued to grind away at his mathematical studies to prepare himself for the Tripos. After nine grueling days of examinations in the bitter cold of January 1879, he

finished third in the competition. His unsuccessful bid for the position of top mathematical wrangler left Pearson bitter and filled with self doubt.

Harboring feelings of failure and resentment and tired of the narrowness of study that he had been confined to in preparing for the Tripos, Pearson used the first year of his fellowship award from the exam to fund a year of study in Germany. There, he pursued the political, philosophical, historical and literary interests he had been developing but had not been free to fully explore (42).

On returning to England, Pearson engaged in a busy academic life which reflected his varied humanistic pursuits. He lectured and wrote on subjects ranging from mathematical physics to German history, politics, and philosophy. From 1880 to 1884, he gave lectures at King's College London on German culture, folklore, and history. His writing and lectures focused primarily on the German Middle Ages and the effect of Martin Luther and the Reformation on German Medieval society (Provine 27).

At the same time he was lecturing about German history and culture, Pearson also returned seriously to the study of mathematics. His interest was in mathematical physics which used the mathematics of known physical phenomena, such as the action of springs and pulleys, to make arguments concerning the activity of unobservable physical phenomena such as the movement of electricity, molecules, etc. (The model for such applications was James Clerk Maxwell—who had scored second in the Tripos behind Routh the year he took the exam.)

Pearson's main source of interest was in the mathematics of the ether, a semi-elastic substance within which all molecules were supposed by some theorists to be suspended and which reacted to their physical movements by telegraphing mechanical

waves across space. In his ether theories describing attraction and repulsion between matter, Pearson took the position that these phenomena could be described by mathematical models based on the harmonious pulsation of atoms within the elastic ether (Porter 187).

Although the content of Pearson's pulsating atom theory was generally dismissed by molecular theorists, his work exhibits important philosophical predispositions which are important to understanding his later arguments about heredity. First, it shows his predisposition to see natural phenomena in terms of existing physical principles or models. In the case of atoms and ether, for example, Pearson believed that their actions could be described from the point of view of the physics of elasticity. Second, it reveals a tendency to invent and develop arguments about natural phenomena at first exclusively from the principles and operations of mathematics and then only later to test through observation whether the mathematical description is supported by empirical evidence. This approach to developing theories of natural phenomena was recognized in Pearson's physical theories and criticized by George Gabriel Stokes, Lucasian Professor of Mathematics at Cambridge. In an 1884 letter to Pearson, Stokes warns him that his calculations of the spectra for polyatomic molecules based on the pulsating atom theory would not hold up when compared to actual observations in spectral analysis (187).

Though his physical mathematical theories never got off the ground, Pearson's knowledge of and skill with mathematics earned him a position in 1884 as professor of applied mathematics in the Department of Mechanics and Applied Mathematics at University College London. In his capacity as an educator, he continued to follow the same philosophical line that he had in his work with physical mathematics, that

theoretical mathematics was a crucial first step in developing a better understanding of actual physical phenomena.

At the same time that Pearson was teaching engineers, he was laboring over the manuscript of *The Grammar of Science* (1892), a text whose purposes were to critique the fundamental goals of science and to reconstruct the rational principles of proper scientific methodology. Predictably Pearsonian in its theory-first approach to the topic, the text introduces some important elements of Pearson's philosophy, particularly his belief that causation does not exist and his opinion about the future of the biological sciences.

Pearson's rejection of causation has important ramifications for his approach to heredity. Primarily, it encouraged him to adopt a probabilistic vision of natural phenomena. This conceptualization explains in part why his vision of heredity was so tightly associated with the probabilistic law of error. Of course he writes,

Scientifically, cause, as originating or enforcing a particular sequence of perceptions, is meaningless—we have no experience of anything that originates or enforces something else. Cause, however, used to make a stage in a routine, is a clear and valuable conception, which throws the idea of cause entirely into the field of sense-impressions, into the sphere where we can reason and reach knowledge. (*The Grammar of Science* [1892], 153)

Cause, in this sense, is a stage in a routine of experience, and not one in a routine of inherent necessity. Causation, says John Stuart Mill, is uniform antecedence, and this definition is perfectly in accord with the scientific concept. (156)

Here Pearson explains that we cannot accept the classical notion of “cause” because we cannot witness one thing causing another at the most fundamental level. All we can observe, he explains is a sequence of stages in a routine. We can infer a relationship between these stages over the course of many observations which tells us whether those events occur in the same sequence over many instances. (In this sense, “cause” can be understood as an observed event which, in most instances, usually precedes another observable event.)

This type of inference is best made using the mathematics of a posteriori probability which allows the user to calculate the probability of a future event, in this case the probability of a particular sequence happening again, based on past observations of this sequence.

Yet the only supposition that we appear to have made is this: that, knowing nothing of nature, routine and anomaly⁹⁰ are to be considered as equally likely to occur. Now, we were not really justified in making even this assumption, for it involves a knowledge we do not possess regarding nature. We use our *experience* of the constitution and action of coins in general to assert that heads and tails are equally probable, but we have no right to assert before experience that, as we

⁹⁰ Used in the same sense here as anomaly, that which is not usual or expected based on previous experience.

know nothing of nature, routine and breach of routine are equally probable. (*The Grammar of Science* (1892), 172-73)

In addition to providing evidence of Pearson's belief in the importance of probability to understanding natural phenomena, the *Grammar of Science* also provides further proof of Pearson's tendency to see all phenomena in terms of theoretical mathematics. Just as he developed a theory of molecular forces and hoped to develop a study of engineering that involved mathematically describable physical forces, Pearson argues in the *Grammar* for a mathematical theory of biology based on the same theoretical mathematics applied to bridges and atoms.

If we look upon biology as a conceptual description of organic phenomena, then nearly all of the statements we have made with regard to physics will serve as canons for determining the validity of biological ideas. In particular, any biological concept will be scientifically valid if it enables us to briefly summarize without internal contradiction any range of our perceptual experience. (395)

Pearson's belief that the approach of the physical sciences could be used to understand biological phenomena was so strong that he concludes the final chapter of *The Grammar*, "The Classification of the Sciences," by proposing a new branch of science in which the methodology of the physical sciences would be applied to the study of biological phenomena.

A branch of science is therefore needed dealing with the application of the laws of inorganic phenomena, or Physics, to the development of organic forms....

Thus just as *Applied Mathematics* links *Abstract Science* to the *Physical Sciences*, so *Bio-physics* attempts to link the *Physical* and the *Biological Sciences* together. (470)

The importance of probability to Pearson's scientific rationale in the *Grammar of Science* and his belief that the techniques and mathematics used in the physical sciences could be applied to the biological sciences provide a theoretical backdrop from which to understand the importance of mathematics, particularly probability, as a source for invention for arguments about heredity.

On the Road to Biometry (1893-96)

Although graphical calculation, scientific theory, probability, and biology may have initially been independent elements within Pearson's spectrum of interests, these elements began to coalesce in the early 1890s into a single unified approach to the study of variation. The coalescence began in 1890 when Pearson was appointed the Gresham Chair of Geometry at Gresham College. As a professor at Gresham, he was responsible for giving twelve one-hour public lectures a year delivered on four consecutive days from Tuesday to Friday during Michaelmas, Easter, and Hilary terms (Magnello, "Karl Pearson's Gresham..." 47). In Pearson's Gresham lectures from 1891 to 1894, he presents two series, one on scientific theory and another on the error curve and graphical

statistics, the latter of which culminates in a geometrical statistical presentation on the process of evolution.

In the second set of lectures on “The Geometry of Statistics” (1891-92), Pearson discussed graphical statistics, the application of geometry to statistical data (Porter 236). In 1891 he discussed death rates and used the graphical statistics presented in the previous lecture to measure variation from the mean. Lecturing the next year on the same subject he introduced the concept of standard deviation. In combination these lectures on graphical statistics, mortality, and standard deviation present the fundamental concepts and techniques which prepared his audience for his Gresham Lectures of 1893-94 which present new mathematical methods for describing bimodal and skewed curves, one example of which had come to his attention through the statistical work of the biologist Frank Weldon. It is this lecture that marks the beginning of Pearson’s foray into biometrics.

Weldon and Pearson

Frank Raphael Weldon (1860-1906) was a zoologist whose primary interest was in the evolution of marine fauna. He had graduated from Cambridge in 1881 and was made a lecturer in invertebrate morphology in 1884. During the earliest period of his research, Weldon worked on developing Darwin’s theory of evolution through the study of morphology and embryology, but he met with no substantial success. After reading Galton’s *Natural Inheritance* in 1889, however, he decided to test the validity of Galton’s

proposition that variation distributes according to the law of errors within a population (Magnello, “Karl Pearson’s Gresham...,” 50).

In May of 1889, he and his wife Florence began collecting measurements of the carapace⁹¹ length of adult female common shrimp *Cragon vulgaris* at Plymouth in Cornwall and Southport in Lancashire.⁹² After corresponding with Galton for help with the mathematics, Weldon was finally able to confirm with his shrimp data that Galton’s argument for the regular distribution of traits was not only true for each separate shrimp population, but also that it was true across all of the populations collected. He published his results under the title, “The Variations Occurring in Certain Decapod Crustaceans. I. *Cragon vulgaris*” in the *Proceedings of the Royal Society* in 1890.

In addition to measuring the *variation* in shrimp, Weldon had also calculated the value of *correlation* between different organs, the degree to which changes in one organ were connected with changes in another. He found a significant correlation across organs in all the samples, meaning that when the measure of one organ changed the measure of the other organ with which it was compared changed as well and to the same degree across all samples. This meant that no matter what the size of the shrimp the relations of their organs remained the same.

Weldon’s work on variation and correlation in shrimp provided further empirical confirmation of Galton’s theory and encouraged him to do more exploration. Weldon’s next subjects were shore crabs *Carcinus moenas*. He collected 23 measurements from two sets of 1000 female crabs, the first from Plymouth sound and then, in the spring of 1892, a second set from female crabs of the same species in the Bay of Naples. Of the 23

⁹¹ The protective back plate of the shrimp.

⁹² He also enlisted the help of a friend in Sheerness near the Mouth of the Thames to collect samples.

characters measured, all but one distributed normally. However, this one exception provided an interesting case which Weldon believed might indicate the presence of a new type in the crab population (Magnello 50).

What he found in his analysis of the frontal measurements of crab shells, was that the curve describing the population was bimodal, double humped, rather than a single hump distributed around one mean which was expected for the data. This discovery engendered excitement in Weldon because a second mean might indicate a new stable type or “sport” as Galton called it. He also, however, believed that it could represent the confluence of two distinct races. With these possibilities in mind, he wrote Galton, “Apart from any arithmetical analysis, I tried to draw inside it two “Curves of Error” whose sum might represent the observed distribution fairly well.... Either Naples is the meeting point of two distinct races or a “sport” is in the process of establishment” (Weldon to Galton 27 Nov. 1892 quoted in Magnello 51).

In order to prove mathematically that either possibility existed, however, it was necessary to isolate the smaller aberrant mean and find the distribution of the characteristics around it. At some point following the spring of 1892, Weldon turned to Pearson for advice about how these calculations might be made. Pearson suggested that a more precise result could be found using a process called curve fitting in which the original values of the data were reevaluated using multiple smaller curves which more precisely followed the contours of the data. Pearson worked on the data for all of the summer of 1893. According to his results, the two humps in the curve divided nicely into two normally distributed curves which he believed revealed the existence of variation as a result of natural selection in the population (52).

From the results of these efforts, Pearson became enthused about the possibility of a purely mathematical approach to the issue of evolutionary variation. His excitement is clear in his notes for his Gresham lectures in 1893 in which he writes,

We are living in an essentially critical period of science, when more exact methods and more sound logic... [are] replacing the old 'scientific gospel'.... For the first time in the history of biology, there is a chance of the science of life becoming an exact, a mathematical science. Men are approaching the question of heredity and evolution from a new standpoint. (quoted in Porter 237-38)

In his November 21 and 22 lectures on probability and the normal curve, Pearson used Weldon's work to discuss the potential usefulness of statistics in solving the problems of evolution, to dispute the ubiquity of the normal distribution in nature, and to discuss his own methods of curve fitting to deal with skewed and bimodal curves. In the text of his November 21 lecture he writes, "Symmetry is by no means universally the case.... The keynote to the most interesting and valuable problems in evolution lies in the non-symmetry of the frequency curves corresponding to the measurements of special organs in animals" (quoted in Magnello 54)). After this statement, he introduces curve fitting for symmetrical and asymmetrical distributions.

In Pearson's writings and Gresham lectures, it is possible to witness the confluence of different strands of his research from graphical calculation to scientific theory and the influence of the statistical work in biology on the direction of these combined interests. In the work of Frank Weldon, Pearson found the data he needed to

illustrate (and, perhaps, even to provide exigence for) his theoretical work with non-normal and asymmetrical curves. It also provided him with an opportunity to make claims about biological processes based on statistical modeling. Though Pearson by no means spent the majority of his time in the latter part of the decade working exclusively on developing biometric theories, he did devote a substantial amount of it on developing mathematical models to describe heredity and evolution. This work, initiated by his contact with Weldon, was instrumental in setting him on a path towards developing the mathematical theory of heredity which is the subject of this investigation.

Oil and Water?

Although Pearson may not have been as excited about the biological ramifications of Weldon's findings as he was about their providing him with material for new methods of graphical statistics, Galton was certainly impressed with Weldon's results. Galton's interest in Weldon's work and Weldon's need for funding to continue it encouraged Galton to call for the formation of a committee to support statistical research on plants and animals. On January 18, 1894, with Galton as the committee chair, and Frank Weldon as secretary, the "Committee for Conducting Statistical Inquiries into the Measurable Characteristics of Plants and Animals" was founded under the charter of the Royal Society for the purpose of funding statistical research regarding questions of evolution and inheritance (Gillham 299).

Initially, Pearson was not a member of the committee; however, his work with Weldon and his zealous support of a mathematical approach to evolution won him a place

on it late in the year of 1896. In the next year, the committee was expanded even further and renamed “The Evolutionary Committee of the Royal Society” at Galton’s request (300).

The reasons for Galton’s 1897 expansion and renaming of the committee appear to be two-fold. On the one hand, it was a way of bringing more traditional practitioners of biology into the group. This move expanded the committee’s breadth and credibility by including, along with biologists doing statistical research, biologists employing traditional methods of qualitative observation and experimentation. On the other hand, the group’s expansion helped to add to the committee researchers whose philosophy of variation was more in line with the position supported by Galton. Whereas members of the committee, particularly Weldon and Pearson, held that variation was predominantly continuous, i.e. that it accreted in small increments over time, Galton believed that it was predominantly the result of discontinuous variation, i.e. that variation was massive and appeared suddenly. (Neither group disagreed with the other about the existence of both types of variation, only about the degree of importance one or the other had to the development of varieties and species.)

One of the new members added was William Bateson (1861-1926), a Cambridge graduate, an early practitioner of the statistical approach to biology, and an ardent promoter of the theory of discontinuous variation. Bateson was a one-time friend of Weldon with whom he had had a falling out because of their constant conflict about the preeminence of continuous or discontinuous variation in speciation. Bateson began to develop his theory of discontinuous variation in 1891 articulating it in an article written by himself and his sister Anna Bateson entitled “On the Variations in Floral Symmetry of

Certain Plants Having Irregular Corollas” in the *Journal of the Linnean Society*. In this article they are cautious not to dismiss the reality and importance of continuous variation; however, they do question the assumption that continuous variation is the only or even primary source for variation in natural forms (Punnett 158-59).

In the next year, the same year that Weldon gathered and examined his data on crabs, Bateson published “Some Cases of Variation in Secondary Sexual Characteristics Statistically Examined,” in the *Proceedings of the Zoological Society* in which he presents statistical evidence for the bimodal distribution of characters in the forceps on earwigs (*Forficula auricularia*) and in the horns on stag beetles (*Xylotrupes gideon*). This work, with results stunningly similar to those published by Weldon in the following year, was, however, different in the sense that it made explicit claims that the bimodal curve was evidence of discontinuous variation. Though Weldon does write to Galton that his data might be evidence of the existence of a “sport,” nowhere does he suggest in his discussion of his data, as Bateson does, that his statistical evidence gives support to a theory of discontinuous variation. Bateson, on the other hand, takes a clear position on the matter.

In the common language of naturalists, the facts of this case [the forceps on earwigs] suggest that there is, for some wholly unknown reason, a dimorphism among males of these earwigs, maintained though all live together....

For the present we are content to recognize that in this case of the earwig there is evidence of a definite and partially discontinuous variation in respect of a secondary sexual character (Quoted in Punnett 197).

Despite any philosophical differences there may have been between Weldon and Bateson about the nature of variation, Pearson saw Weldon's crab curves and Bateson's earwig curves as statistically similar and included both of them as examples of bimodal curves in the printed abstract of his paper "Mathematical Contributions to the Theory of Evolution" (1893). He writes, "Such abnormal curves arise particularly in biological measurements; they have been found by Professor Weldon, for the measurement of a particular organ in crabs, by Mr. Thomson for prawns, by Mr. Bateson for earwigs" (Pearson, "Mathematical Contributions," 329-30).

Though their data might have distributed in the same way, Bateson makes it very clear in his major work, *Materials for the Study of Variation* (1894), that he supports a strict theory of discontinuous variation. In *Materials*, Bateson catalogues facts about variation in specific organisms and speculates about the nature of variation based on these collections of facts. He argues that the evidence supports his claims 1) that discontinuous variation exists, 2) that this variation has its source within the organism, and 3) that this type of variation is not a product of natural selection.⁹³ He writes, "The discontinuity, of which species is an expression, has its origin not in the environment nor in any phenomenon of adaptation, but in the intrinsic nature of organisms themselves, manifested in the original discontinuity of variation" (*Materials* 567).

Once he has sufficiently argued for the existence of discontinuous variation and its likely role in speciation, Bateson concludes the text with a discussion about the methods whereby these issues might be further explored and clarified. He argues in favor

⁹³ Though Bateson makes the case that discontinuous variation exists and that it plays the seminal role in speciation, he does not dismiss the existence of continuous variation.

of a continued program of gathering statistical data on discontinuous variations; however, he explains that this program should not be purely observational but should rely also on experimental breeding to direct or limit the scope of the data set: “The only way in which we may hope to get the truth is by the organization of systematic experiments in breeding, a class of research that calls perhaps for more patience and more resources than any other form of biological inquiry” (574).

As an indirect critique of Bateson’s book, Weldon published “Remarks on Variations in Plants and Animals” in the *Proceedings of the Royal Society*. In this paper, Weldon makes two things clear about his position on variation. First, he explains that the only way evolution can be profitably examined is through the use of statistics. This examination includes knowing the distribution of values for a particular trait around the mean as well as the frequency of the birth or death rate of organisms within the population for different values within this distribution. Second, understanding evolution does not require an understanding of heredity. Weldon believes it is not necessary, or perhaps even possible, to know how organisms vary between generations. All we need to know is the magnitude of a particular character for each generation and the difference between magnitudes across generations to show the rate of variation.

It is to be observed that numerical data... contain all the information necessary for a knowledge of the direction and rate of evolution. Knowing that a given deviation from the mean character is associated with a greater or less percentage death rate in the animals possessing it, the importance of such a deviation can be estimated without the necessity of inquiring how that increase or decrease in the

death rate was brought about.... In the same way a theory of the mechanism of heredity is not necessary in order to measure the abnormality of an offspring associated with a given parental abnormality. (“Remarks” 381)

In these words, Weldon clarifies the difference between his and Bateson’s position. Bateson believes that it is possible to find the source of variation within the organism. Weldon feels that all the necessary information about variation can be gathered statistically from the outward measure of characters. Because the source of variation for Bateson is internal, knowledge of heredity from experimental breeding provides the best route for understanding variation. For Weldon, the examination of heredity is a waste of time. He does not need to know the cause of variation to examine its effects on differential survival.

The differences between these two biologists are important to understanding the position taken by Pearson in his hereditary arguments and the arguments presented against his position. By the time Pearson makes his case for inheritance, he has been working with Weldon closely for ten years; therefore, it is no surprise that his mathematical model of heredity is informed by Weldon’s biological assumptions. In fact, Pearson’s hereditary argument is actually a defense of Weldon’s model in that it provides a theory of heredity which characterizes the phenomenon as simple and mechanical and, therefore, unworthy of detailed scrutiny. This exigence explains in part why Pearson chooses the error curve and its underlying conditions as the main source of argumentation.

The Biometrik Response

Given the animosity that developed between Weldon and Bateson over their biological theories in the years following the formation of the “Committee for Conducting Statistical Inquiries into the Measurable Characteristics of Plants and Animals,” it was no surprise that despite Galton’s good intentions to bring traditional biologists, mathematicians, and quantitatively inclined biologists together, the resulting combination was a disastrous mix. What was initially a difficult marriage between the two groups became by 1899 a completely untenable situation. According to historian William Provine, under the impression that Galton, the lynch pin that was holding the group together, had changed his position on the importance of the study of heredity mechanisms by accepting Pearson’s revision of his law of ancestral heredity, Bateson, through some skillful maneuvering, attempted to turn the interests of the committee towards a program of plant and animal breeding (Provine 54).

Realizing that the committee had become untenable, Galton, Pearson, and Weldon all resigned in January of 1900 in the hopes that the committee would collapse. Through some shrewd political wrangling, however, Bateson garnered enough support to keep it from folding. As a result, the committee was completely dominated by Bateson and his supporters much to the chagrin of Pearson who later wrote that their “capture of the committee was skillful and entirely successful” (Gillham 307).

Angered by Bateson’s success in getting control over the evolutionary committee and turning it towards breeding, a of type investigation that Weldon felt was profitless, Weldon set out, with the firm backing of Pearson and auxiliary support from Galton, to

establish a conduit for their descriptive statistical investigations of biological phenomena. To create a conduit for their biological work, they founded the journal *Biometrika* in 1901. Its purpose according to Pearson's opening editorial, "The Scope of *Biometrika*," was to, "serve as a means not only of collecting under one title biological data of a kind not systematically collected or published in any other periodical, but also of spreading a knowledge of such statistical theory as may be requisite for their scientific treatment" ("The Scope of *Biometrika*" 1).

Although the opening lines of Pearson's editorial seem to represent some common-sense goals of gathering data and developing a statistical approach to biological questions, further reading reveals that *Biometrika* was also intended to support a more radical platform for the development of a new mathematical biology guided exclusively in its investigations of evolution and heredity by the principles and practices of mathematics, especially statistics and probability. In the second section of the editorial entitled "The Spirit of *Biometrika*," Pearson expresses his views that biometrics must be focused solely on a mathematical solution to the questions of evolution, variation, and heredity.

Whatever views we hold on selection, inheritance, or fertility, we must ultimately turn to the mathematics of large numbers, to the theory of mass phenomena, to interpret safely our observations. As we cannot follow the growth of nations without statistics of birth, death, duration of life, marriage and fertility, so it is impossible to follow the changes of any type of life without its vital statistics. The

evolutionist has to become in the widest sense of the words a registrar-general for all forms of life. (“The Spirit of *Biometrika*” 3)

In his mission to create an exclusively statistical domain in the study of evolution, variation, and heredity, Pearson hoped to reach two different audiences. He makes the dual nature of his audience clear in his description of the statistical scientific enterprise which included both biologists and mathematicians.

The biologist may find in our pages algebraic analysis which may repel him. We would still ask his attention for the general conclusions and for the formulae reached by the mathematician. The biologist will find that they frequently suggest observations and experiments which he alone is in position to undertake satisfactorily. We shall aid the more arithmetical part of his work by diagrams and numerical tables wherever it seems possible. In this manner we hope that *Biometrika* will provide for both branches of science; that it will not only publish valuable biometric and statistical researches, but serve as a storehouse of unsolved problems for both unemployed biologist and mathematician. (5-6)

As in his earlier days working with engineers and with physical mathematics, this characterization of the relationship between the mathematician and the scientist suggests a hierarchical relationship with mathematics and the mathematician at the top. In this scheme, the mathematician’s job is to come up with formulae and to reach general conclusions which then suggest the type of data to be collected and the subject and

manner of experimentation to be carried out by the biologist—which is exactly what the rhetorical topoi do. According to this biometric vision, science becomes the process of developing and testing mathematical theory through the collection of data.

Although *Biometrika* had been founded as a safe place for Weldon and Pearson to present their statistical-biological work, it was not the only outlet for Pearson's efforts. In fact, Pearson had been publishing mathematical papers based on biological examples in the series "Mathematical Contributions to the Theory of Evolution" in section A of *The Philosophical Transactions of the Royal Society*, the section devoted to mathematics, since 1893 and continued to publish there until 1904. At first, the titles of these papers reveal that the primary topics of discussion were either the new statistical tools used in biological investigations or new statistical procedures that could be developed to analyze statistical problems presented by biological data. In 1895, however, the goal of some of these papers began to change including titles suggesting that they were written to establish biological principles to support the mathematical models applied to biological phenomena. The appearance of these new titles seems to be a response to a new conflict within the statistical committee itself between Galton and Pearson/Weldon over whether the continuous accretion of variation was sufficient to supply the larger variations which could be acted upon by natural selection to create new varieties and species (Provine 35).

With the rediscovery of Mendel's work at the turn of the century and the development of Hugo de Vries' theory of mutation in the same year, the pressure on the Darwinians to defend their theory of variation by small steps mounted. In particular, it was necessary for them to prove that most variation was both regular and small. It had to be small; otherwise, it would be considered discontinuous. It also had to be regular. In

order to measure and compare variation to show the direction and magnitude of evolution using statistics, the variation measured had to be uniform across the population and change predictably from one generation to the next. If the variation was too haphazard either within or between generations, then no available mathematical modeling could apply.

In addition, Weldon and Pearson had to defend their position that their model was sufficient to explain evolution. This position was challenged by both Bateson and Galton who made the case that small variations would be swamped or blended away in subsequent breeding and, therefore, would never accumulate sufficiently to be acted upon by evolution. Only radical large changes created by the appearance of stable ‘sports’ could serve as the material for selection.

Realizing that their mathematical model of heredity was being seriously challenged, both Weldon and Pearson began to shore up their defenses. Weldon took on Mendel’s work in his paper “Mendel’s Laws of Alternative inheritance in Peas” making the case that Mendel’s results did not take into account the inherited contributions of remote ancestry. Pearson, on the other hand, took on aspects of Galton’s work in part nine of his series “Mathematical Contributions to the Theory of Evolution” in an article entitled “On the Principle of Homotyposis and its relation to Heredity.” In this article, Pearson proposes a theory of heredity which challenges Galton’s belief that heredity is a process of selection and combination which takes place inside the organism. Though both Weldon’s and Pearson’s arguments are important to the debate, Pearson’s is of greater interest here because it is directly involved with supporting the mathematical model developed by himself and Weldon using arguments from the special topics of the error

curve. Also, Pearson's communications with Bateson regarding this publication reveal the importance of mathematics to establishing the credibility of scientific arguments.

Homotyposis and Heredity

In Pearson's homotyposis paper, he proposes a model of heredity in which variation in organisms is the result of external and random phenomena which alter the reproductive material when it is produced. This particular vision of heredity supports his and Weldon's statistical Darwinian model in which variations had to be small and regular. At the same time, it challenged Galton's conception of variation as an internal process driven by selection and combination.

In order to make the case for the minuteness and regularity of variation, Pearson argues for the existence of hypothetical organs called *homotypes* or *undifferentiated like organs* which by definition vary regularly and minutely within a generation and from one generation to the next according to the law of error. Pearson's choice of the phrase "undifferentiated like organ" to describe this central phenomenon is unfortunate in that it appears at first glance either contradictory or unnecessarily redundant. Though Pearson himself is, "fully conscious of the *verbal* difficulty of the phrase," he nonetheless believes that the distinction between differentiated and undifferentiated is mathematically clear (Pearson, "On the Principle of Homotyposis," 287). Essentially, the difference between these two phenomena is in their source and graphical distribution. An undifferentiated population is homogeneous and distributes according to the law of errors symmetrically around a single mean. A differentiated population, on the other hand, is by

definition heterogeneous and does not distribute normally around a single mean, like Weldon's crabs. Because it is made up of two or more populations of types with different means, its values will distribute multimodally with humps around each of the means of the population types mixed in the sample (287).

According to this explanation two things can be said about undifferentiated like organs. First, because they distribute according to the law of errors, their variation is the result of a large number of small, unknown influences. Second, we know that they are all essentially expressions of some true value. The former point is explicitly stated by Pearson when he writes, "variability in 'undifferentiated like organs' is not to be associated with any one or two dominating factors which could be isolated, it is due to that combination of many small causes, inherent and environmental, which leads to... a homogeneous chance distribution" (287). The latter point is made in his use of the word "like" in the phrase "undifferentiated like organs." What he means by this term is that the organ being described is produced from a single non-variable mold or true value. He makes this point when he writes:

I shall call undifferentiated like organs *homotypes* indicating that they are types produced by the same mold or individual. Thus two leaves of the same tree or two blood corpuscles from the same frog are *homotypes*; their resemblance will be homotypic, and the character by which their resemblance is quantitatively measured will be the homotypic character. (294)

Pearson's model of similarity and variation in terms of replication and error meets the criteria for supporting the mathematical practices and models that he and Weldon developed to make their case for continuous variation. First, the process is regular in the sense that a feature is struck from the same mold and that the variation which affects it can be described by an error distribution. This allows the results to be mathematically rendered and makes variation subject to the principles of the law of error. Second, it keeps variation small which is a necessary criterion of Darwin's model.

Once Pearson has argued for the existence of organs whose features meet the specifications of his and Weldon's model variation, he goes on to make the case that heredity and variation can be understood in terms of the variations found in like "organs" in a single individual. In order to make the connection between undifferentiated like organs and heredity/variation, he first argues that the sperm and ovum are undifferentiated like organs: "But turning to the process of reproduction, the offspring depend on the parental germs, and it would thus seem that the degree of resemblance between offspring must depend on the variability of the sperm cells which may be fairly considered "undifferentiated like organs" ("On the Principle of Homotyposis" 287-88). Then he argues that if the sperm and ovum are considered undifferentiated like organs, the process by which they vary must be the same process by which undifferentiated like organs vary: "Inheritance will not be a peculiar feature of reproductive cells.... If this view be correct, variability is not a peculiarity of sexual reproduction, it is something peculiar to the production of undifferentiated like organs in the individual, and the problems of heredity must largely turn on how the resemblance between such organs is modified" (288).

Based on Pearson's assumptions that sperm and ovum are undifferentiated like organs and that the process of their production is the same as the production of undifferentiated like organs, it is possible to understand the similarities and differences between Galton's model of heredity and Pearson's. In one important way, both Galton and Pearson's models are the same. They both hold that chance plays an important role in variation and, as a result, that the error curve can be used to describe it. The difference between the two, however, is how chance influences hereditary outcomes. For Galton a large number of unknown causes influence the selection of traits from a pool of available features, the influence of this selection being evident most markedly in the normal distribution of traits which blend.

Unlike Galton, however, Pearson holds that the action of chance on reproduction does not operate on the selection of internal characters during fertilization. Instead, chance acts in the production of undifferentiated like organs, a process which can be analogized to a factory production line. Like a factory, the body of the male or the female produces either sperm or ovum from the same mold. During the process of production, however, many small unknown influences affect the replication creating slight variations in the sperm and ovum produced. These variations are then collectively passed on to offspring.⁹⁴ As a result, the variability observable in the phenotype of the offspring has its source not in the characters inherent in the sperm or ovum or their combination, but rather in errors during the replication of these undifferentiated like organs.

Once Pearson has argued that the sperm and the ovum are undifferentiated like organs and that the hereditary process is the reproduction of undifferentiated like organs,

⁹⁴ Pearson assumes that the contributions of each of the parents are equal and blend together during most instances of reproduction.

he has to provide sufficient evidence to support his case. He believes that if he can show both mathematically and empirically that the correlation value, i.e. the degree of relatedness, between undifferentiated like organs is the same value as the correlation value between offspring, this would support that they were produced by the same process.

Pearson begins by mathematically proving that, “*The correlation between brothers... [is] equal to the mean correlation between the undifferentiated like organs put forth by an individual*” (“On the Principle of Homotyposis” 291). He offers a mathematical proof for this relationship in the second section of the article entitled “Theoretical Discussion of the Relation of Fraternal Correlation to the Correlation of Undifferentiated like Organs” with equations for variation, correlation, mean and deviation based on the principles of the error curve.⁹⁵

Once he has deduced mathematically that the average of the combined correlation values of sperm and ovum and the correlation of offspring should be equal, Pearson provides observational evidence from undifferentiated like organs to support his case. In his opening statement of the third section of the homotyposis paper, “On the Variability and Correlation of Undifferentiated Like Organs in the Individual,” he explains the rationale behind his choice of evidence.

What I have endeavored to do is to take as wide a range of different organs as possible in different types of life and trust to the bulk of my statistics to give me a substantially accurate value of ρ [correlation of undifferentiated like organs] to

⁹⁵ See Appendix M

compare to the values of **R** [fraternal correlation] we have determined on other occasions. (292)

True to his word that he is going to give evidence concerning, “as wide a range of organs as possible,” Pearson provides correlation values from undifferentiated like organs such as leaf veins, mushroom gills, and Shirley poppy seeds. Using leaf veins as an example, he explains his statistical method for calculating the correlation. In conducting this type of analysis, he gathers 26 leaves each from 100 trees in the same district at a given height. He then counts the number of veins in each leaf. Next, he places these values on a correlation table pairing each leaf with every other leaf to generate a total of 325 entries ($\frac{1}{2}(26 \times 25)$). In order to maintain symmetry in the correlation values, he also creates a second table in which the order of comparison is switched. This results in a total of 650 entries for each tree and 65,000 values for 100 trees. Once all the values are gathered, he calculates the mean and standard deviation of the sample which allows him to generate a correlation value for the data (292-93, 295). Finally, he compares the mean values of correlations to the fraternal correlation **R**= .4 as it is described by his version of Galton’s law of ancestral heredity.

Pearson’s ability to make his case using a wide variety of correlative values from different types of undifferentiated like organs is only possible because of his assumption that the difference in the production of these organs in all organisms is uniform and small, i.e., it is the result of regular error precipitated by a large number of small unknowable causes. He confesses the importance of this feature of the error curve in his mathematical section when he writes,

Now if the degree of resemblance between undifferentiated like organs vary largely from organ to organ we could proceed no further with this investigation. We should have shown that there was a relation between, namely (iii)⁹⁶ between the **R** [fraternal correlation] and the ρ 's [correlations of undifferentiated like organs] but as the characters in the spermatozoa are unknown and unmeasurable we could not proceed further. On the other hand, if the correlation between undifferentiated like organs is approximately constant, than we may replace correlation coefficients... with their *mean* value ρ , say. ("On the Principle of Homotyposis" 290)

Here Pearson admits that without assuming that variation is regular and small there would be no way of making a general case about the process in heredity in the large population using correlation values measured in individual instances. He argues, however, if variability is small and we can find a regular correlation in measurable features, such as the veins in tree leaves in a sample of trees, this value can then be used along with the average values of other trees to determine the correlation value of leaf veins in the whole population. In addition, because correlation is measured from the mean of a sample in units of its own standard deviation, it is a standard unit of measure (Chatterjee 297). As a result, it can be used to compare variables as diverse as mushroom gills and leaf veins. Because correlation allows for both an inter- and intra-specific

⁹⁶ See page 290 of the section of Pearson's paper entitled "Theoretical Discussion of the Relation of Fraternal Correlation to the Correlation of Undifferentiated like Organs" provided in Appendix M.

comparison, Pearson believes that if a mean value of correlation of all undifferentiated like organs for all forms of life exists, he will be able to find it using his method.

Though Pearson hopes that his model of heredity based on the law of error will allow heredity to be described within a circumscribed quantitative range using correlative values, his results suggest that this model of heredity is problematic. He begins the results section of his paper by making an ethical appeal to his audience to accept his results despite their imperfections based on the claim that he has been honest in reporting his findings and has not tried to obscure data which he might have left out to make his case stronger.

In summing up my results and comparing them with those obtained from fraternal correlation by my coworkers and myself I felt some difficulty. If I made selection of what I considered the best homotypic correlation series and the best fraternal correlation, I might well lay myself open to the charge of selecting statistics with a view to the demonstration of a theoretical law laid down before hand. (“On the Principle of Homotyposis” 355)

After making this appeal, Pearson reports all of his figures.⁹⁷ Despite this show of integrity, however, he does not actually use all of his results in calculating the mean correlation for undifferentiated like organs. He chooses to leave out the data gathered from the measurements of correlation in mushroom gills and ivy leaves because he believes these measurements are unduly influenced by “correlation due to stages of

⁹⁷ See the chart labeled “General Results for Homotypic Correlation” in Appendix N “

growth” (356). This elimination is acceptable in Pearson’s opinion because such growth would be considered a differentiating influence.

A closer inspection of the summary of values on the chart, however, reveals that Pearson’s elimination of these values, at least the values for ivy, seems to be motivated exactly by a desire to fit the data rather than for any valid exception of outlying data. The correlation values he presents for the length of ivy leaves, their breadths, and the combination of their lengths and breadths are respectively .56, .53, and .51. A glance at the values on the acceptable portions of the chart reveal that there are no less than 8 of a total of 22 acceptable measures which actually exceed the highest value presented by the ivy and a total of 11 that are larger than the lowest value in the ivy measurements. If the ivy results based on their value are in fact aberrant, then so are half of the results presented as acceptable.

Basing the exclusion of the results on factors of growth seems even more treacherous for Pearson who admits that he is not a trained botanist and that quantitative data are the only acceptable means of assessing inheritance. This leaves his argument caught up in a contradiction whereby it accepts data as the arbiter of truth but then rejects quantitative results based on supposed qualitative conditions.

In addition to irregularities evident in Pearson’s eliminating certain values from his calculation of overall correlation, there is also an unusual tendency on his part to admit data which might reasonably be eliminated based on their values falling far outside the normal range. If we take Pearson’s own calculation of the mean for the correlations in his research, .4570, and the highest correlation value he has accepted into his normal distribution, .6313, we can assume that he believes that reasonable values fall within

.1743 of the mean. If this is the extent of the accepted values for the lower limit, then for the sake of symmetry of outlying values in a binomial distribution there should be a similar limit below the norm. This would mean that any values below .2827 should be eliminated as well. A brief inspection of the lower values on Pearson's chart, however, reveals that he has not applied this practice to his data. In fact four or more than 1/3 of the data in his chart, fall below this value, yet they are still included in the final computation.

The combined result of leaving out values which fall above the mean but below the upper limits established for the data and including data whose values fall below the reasonable limits is that the mean value of the distribution is artificially lowered. A recalculation of the results, which extends the upper limits of acceptable data to the same degree as the lower limits are extended and includes all of the data which are excluded by Pearson but fall well within the limits of the distribution, results in a shift of the mean upwards. The new upper limit based on a recalculated spread of the acceptable lower values (.2837) from the mean is .7407. This new limit includes all of the unacceptable data with the exception of the length of mushroom gills. By including all the data within the new acceptable range, the mean value of correlation moves from .4570, which already differs from his original hypothesis of .4, to .4826 an even greater difference.

Further evidence that Pearson is massaging data to get a better fit with his theory can be found in his reevaluation of the fraternal correlation⁹⁸ to a value higher than his own original estimation at the beginning of the paper. In the beginning of the text, Pearson clearly states that the value of the fraternal correlation is .4. In the results section, however, he gives a different value for this correlation, .4479. This new calculation is made from a combination of statistics taken from plants and animals as well as humans.

⁹⁸ See the "General Results for Fraternal Correlation" chart in Appendix N

His calculations include such correlations as the coat color in basset hounds, the same data Galton used in developing his Law of Ancestral Heredity, and horses, whereas the original calculation is based solely on the statistics gathered from human brothers. The correlation statistics for plants and animals are all substantially higher than those in humans; the highest correlation value for humans is .5424, a full .15 lower than the highest correlation value for *Daphnia* at .6934, pulling the average value of the correlation considerably higher.

Ignoring these obvious discrepancies in data selection, Pearson argues that the similarity between this new value for fraternal correlation .4479 and his average correlation for undifferentiated like organs .4570 is close and seems to provide proof for his conclusions 1) that sperm and ovum are undifferentiated like organs and 2) that reproduction is a case of the production of undifferentiated like organs. In other words, progeny differ from each other just as leaves on the same tree do.

I do not propose to lay great stress on what at first sight might look like a most conclusive equality between the mean values of homotypic and fraternal correlations—within the limits of probable errors .4479 and .4570 are indeed equal. I am quite aware that a few further series added to either the homotypic or fraternal results might modify to some extent this equality. But what I would ask the reader to do is to examine the two tables side by side, to note how the first and last several results of both may be fairly held to be subject to quite definite modifying factors, and then to consider whether there is not very substantial evidence gathered from a fairly wide range of characters in nearly as wide a range

of species to show that homotypic and fraternal correlation fluctuate about a mean value between .4 and .5. (“On the Principle of Homotyposis” 358)

Despite all of his efforts to minimize the difficulties inherent in trying to reduce biological qualities to exact quantities within his mathematical model, Pearson is forced to recognize the formidableness of the obstacles facing his hereditary theory. Following his efforts to legitimize his conclusion he writes,

I do realize that it is extremely difficult with the complex system of factors influencing living forms to reduce our conditions to that theoretically perfect state in which we shall measure solely the factor we are investigating.... In the first place the theoretical conception of undifferentiated like organs is very hard to realize practically.... Secondly, the environmental factor comes into play. It is difficult to obtain a hundred individuals with like environment; soil, position with regards to other growths, sunlight, insect life, etc.... Thirdly, the difficulty that ensuring all individuals are of the same age or in the same stage of development, is very great. (358)

Undaunted, Pearson remains confident that the principles of the error curve can be usefully employed to describe the action of heredity. He closes his paper by making an argument from the law of large numbers which holds that the more data which are gathered on a phenomenon, the closer the researcher can get to its true value. Within a large collection of data, Pearson argues, the differentiating factors which would skew the

true value of the sample either upwards or downwards would cancel each other, leaving only the true value for the “bathmic influence,” i.e. the influence of regular error on the undifferentiated like organ.

If homotyposis had a practically constant value throughout nature, I should only expect this value to be ascertained as the result of the average of many series in which the opposing factors of differentiation... may more or less counteract each other. In this manner we may approach to a fair appreciation of the bathmic influence of individuality [the value of the set of many small unknown causes that effects the production of sperm and ovum in individuals] in the production of undifferentiated like organs. (“On the Principle of Homotyposis” 359)

In his concluding remarks in the text, Pearson expresses his confidence that his evidence and arguments have sufficient credibility to support his conclusions 1) that the value of correlation found for the production of different organic homotypes, leaves, mushroom gills, etc. was remarkably similar to the correlation values found for different physical features in brothers and 2) that, as a result of this similarity, it is logical to conclude that heredity is an example of the process of homotyposis. Because homotyposis produces small and regular variability under the same conditions assumed by the mathematical law of errors, the statistical modeling that Pearson and Weldon have been using and promoting as the only method for studying the magnitude and direction of evolution is justified. In addition, homotyposis supports the veracity of a Darwinian model of evolution in which the major variations that are the basis of specific formation

through natural selection are developed by the accretion of small variations in offspring. Finally, it refutes Galton's hereditary model in which variation finds its source within the organism and is the result of the random selection of phenotypic features from a genotypic pool of traits.

At first blush, Pearson's argument appears to be a triumph of theoretical mathematical reasoning in making biological arguments. It begins with a hypothesis proved deductively from mathematical principles and incorporates the mass collection of empirical data to test the validity of his conclusions. A closer inspection of his methods for selecting data to make his calculations and of the challenges surrounding the identification of pure homotyposis (i.e. a process of replication/variation unaffected by factors which might increase or decrease the degree of variation/similarity beyond the limits allowed by chance influences) reveals, however, that there are major weaknesses in his model.

Bateson's Critique

In the fall of 1900,⁹⁹ the same year that Weldon, Pearson, and Galton left the evolutionary committee, Pearson submitted and presented his Homotyposis paper to the Royal Society. A few months later, in January and February of 1901, William Bateson submitted and read a critique of Pearson's ideas in front of the same body. The fact that Bateson responded to Pearson's attempts to defend his mathematical model of Darwinian variation comes as little surprise given Pearson's close relationship with Weldon and the

⁹⁹ In the spring of 1900, some time after May 8, Bateson became acquainted with Mendel's "Experiments in Plant Hybridization."

friction between Bateson and Weldon over the priority of continuous or discontinuous variation in supplying the material for specific variation. What may come as a surprise, however, is that up to the publication of Pearson's article and even in his critique, Bateson was actively trying to recruit Pearson to abandon Weldon's theory of continuous variation. Why would Bateson try to recruit his arch enemy's right hand man? I contend that his actions were motivated by his belief in the power of mathematics to make or break scientific arguments.

This position is not without controversy, however, because some historical evidence describes the debate as antagonism between Bateson's biological method and the mathematical methods developed by the biometricians. A commonly used phrase around the time of the disputes which developed between Bateson and Weldon/Pearson, for example, was the "Biometrician-Mendelian controversy." This antithesis suggests that the definitive difference between the two sides in the debate was that the biometricians were mathematical while the Mendelians were non-mathematical in their approach to heredity, a position disputed in this chapter and in the chapter on Mendel. Evidence that biologists at the time of the debate split the groups along the lines of mathematics and biology is present in a 1910 book review by George Shull of Wilhelm Johannsen's (1857-1927) *Heredity as an Exact Science*.

An essential part of the reaction [which occurred at the end of the nineteenth century against speculative theories of evolution and heredity], the development of mathematical methods by Professor Karl Pearson and a few others promised for a time the key to the riddles of evolution. This movement, led by a pure

mathematician, developed a series of beautiful methods for the mathematical analysis of data and the comparison of variations....

Parallel with this movement toward the use of mathematically precise methods, there has been a rapidly increasing utilization of pedigree culture or genetic methods... in Mendelian investigations of heredity which methods lay stress chiefly on biological methods of analysis. The antagonism of active workers with biometric and genetic methods in the study of heredity scarcely permitted it to be hoped that a work might soon appear... utilizing the results gained by both biometry and genetics. (Shull 226)

Shull's division of the two groups into the biologists and the mathematicians was most likely encouraged by comments made by Pearson who, being the only professional mathematician engaged in debates with a group of biologists, often took criticism of his qualitative biological assumptions as an attack on his novel mathematical approach by backward-looking biologists. For example, Pearson wrote Galton after the first meeting of the new evolutionary committee which included Bateson to complain,

The committee you have got together is entirely unsuited.... [It] contains far too many of the old biological type, and is far too unconscious of the fact that the solutions to these problems are in the first place statistical, and in the second place statistical, and only in the third place biological. (quoted in Provine 51)

This statement reveals that Pearson believed that there was a division among members of the group, which now included Bateson, along the lines of traditional qualitative biologists and modern mathematical biologists. Statements such as this one and others like it may explain why modern historians rely on the division between biology and mathematics in the two groups to explain why the Evolutionary Committee of the Royal Society became disorganized: “the new members showed little sympathy to the biometrical approach, and in some cases antagonism” (Provine 51). The split is also used as the motivation for the formation of *Biometrika*: “*Biometrika*, established in 1901, became the official organ of the new statistics the standard bearer of a quantitative program in opposition to the biological mainstream” (Porter 269).

Although these characterizations of a split between the biological and the mathematical camps seem to be justified by certain historical evidence, other evidence suggests that an alternative interpretation of the divisions between the two groups can be made in which Bateson is a supporter and not a detractor of the biometrical approach. An historical review of Bateson’s work, for example, reveals that there is ample evidence that he can be fairly grouped among biologists who relied on mathematics as a source for evidence and invention in his arguments. His use, for example, of the law of errors and the error curve to draw conclusions about the dimorphism in stag beetles and earwigs two years before he published his work on dimorphism in shore crabs suggests that, if anything, Bateson had beaten the biometricians to the punch.

In addition to the evidence that Bateson was not opposed to biological arguments which drew support and inspiration from mathematics, there is also evidence in Bateson’s history with the biometricians that his primary critique of their mathematical solutions to

biological problems was not that they were mathematical, but rather that they were only, as Pearson contends, in the third place biological. He had complained, for example, that that the work done by Weldon on the carapaces of *Carcinus moenas* in measuring variation did not take into consideration the fact that he had not measured crabs in the same stage of molting. As a result, the distribution of the shell characteristics he was measuring would not be accurate and, therefore, could not provide evidence for the action of continuous variation in evolution (Provine 49).

In the events surrounding the publication of Pearson's paper and Bateson's response, we can actually see the same high regard for mathematics in investigating biology and the same criticism of the biometrician's disregard for biological phenomena that is evident in Bateson's previous work and critique. In the winter or spring of 1901, conflict arose between Pearson and the Royal Society when, without Bateson's permission, the manuscript of his critique of Pearson's paper was distributed to referees before the referees had received a complete manuscript of Pearson's work. Needless to say, Pearson was angered by this unorthodox procedure and saw it as more evidence of discrimination in the biological establishment against his attempts to develop a mathematical biology (62).

When Bateson found out what had happened, he immediately withdrew his critique until Pearson's paper was published. He also promptly wrote him a letter apologizing for the incident. Bateson's response soothed Pearson's ire to some degree and initiated an interesting exchange of communication between the two over the course of the next year (62-63).

The proof that Bateson saw mathematics as an important source for biological arguments is made clear in his attempts during his correspondence with Pearson from 1901 to 1902 to recruit Pearson and his mathematical talents away from Weldon. In October of 1901, Bateson sent a translation of Mendel's paper to Pearson in the hopes that, after reading the text, the mathematician might be persuaded to accept Mendel's theory of inheritance. Though Pearson read it, in his reply he "expressed skepticism about the general applicability of Mendelian inheritance" (Provine 63).

Even though Pearson was not receptive to these initial efforts, Bateson remained hopeful that he might yet persuade Pearson to join forces with him. With the publication of Weldon's critique of Mendel in *Biometrika* (January, 1902), however, Bateson realized that the level of enmity developing between the two camps was such that his window of opportunity for recruiting Pearson was almost gone. In a last ditch effort, he wrote Pearson on February 13, 1902, exhorting him to reconsider.

I respect you as an honest man and perhaps the ablest and hardest worker I have met, and I am determined not to take up a quarrel with you if I can help it....

There has probably been no discovery made in theoretical biology that we can remember which approaches Mendel's in magnitude, and the consequences that it leads to. This is not a matter of opinion but certain. You have worked well in the same field and if through any fault of mine you were to be permanently alienated from the work that is coming, I should always regret it. With Weldon it is different. He is a naturalist. He goes in with his eyes open. (63)

Ultimately, Bateson's efforts failed. In his response, Pearson tried to maintain an appearance of rational balance by writing that though his heart was with Weldon his head was still open to various positions. Despite this posture of fairness, Pearson made it known in no uncertain terms that he was firmly behind Weldon's philosophy of continuous variation. Shortly after this exchange, Bateson's critique of Pearson's paper "Heredity, Differentiation, and other Conceptions of Biology: A Consideration of Professor Karl Pearson's Paper 'On the Principle of Homotyposis'" (1902) was published as well as a scathing attack on Weldon's critique of Mendel titled *Mendel's Principles of Heredity: A Defense*. Any chance of détente between the two men vanished.

Though 1902 marked the end of friendly communications between Bateson and Pearson, it was not because Bateson was an ardent foe of biometrics. His interest in recruiting Pearson's talents to help further develop the theory of Mendelian inheritance suggests that, though Bateson was at the core a biologist dedicated to experimental breeding as a way of unlocking the mysteries of heredity, he was also keenly aware of the power of mathematics as both a tool and as a source for making biological arguments.

The respect that Bateson shows for the use of mathematics to make biological arguments in his communications with Pearson is similarly evident in his critique of the mathematician's homotyposis paper. In this critique, Bateson does not attack Pearson for making an argument about the nature of heredity using mathematics; instead, he attacks the validity of the biological model which Pearson's mathematical arguments are supporting.

In the beginning of the text, Bateson seems to continue his overtures to Pearson by making the case that they both share similar interests. He commends the

mathematician for adopting a philosophical position that is near his own and draws parallels between Pearson's concept of homotyposis and his own work in *Materials for the Study of Variation* (1894) in which he argues that variation has its source in symmetry and repetition (*Materials for the Study of Variation* 242).

At the outset I wish to express the conviction that the leading idea which inspired and runs through the work is a true one. Professor Pearson suggests that the relationship and likeness between two brothers is an expression of the same phenomenon as the relationship and likeness between two leaves on the same tree, between the scales on a moth's wing, the petals of a flower and between repeated parts generally....

This idea first came to me—as it has perhaps to others—when I was studying the phenomena of variation in meristic series,¹⁰⁰ and in writing on that subject I introduced an outline of the conceptions involved. On that occasion I ventured to carry this reasoning a step further... and to suggest that *the resemblance which we call heredity may be a special case of the phenomenon of symmetry*. (Bateson, "Heredity, Differentiation," 194)

Following a few complimentary words, Bateson lays out his critique of Pearson. He argues that there are essentially two major problems with his theory of homotyposis. First, given the current qualitative biological knowledge of organisms, it seems unlikely

¹⁰⁰ Meristic variation refers to changes in the number and geometric organization of the physical features of an organism such as, for example, changes in the number of petals or their distribution on a particular species or variety of flower. Bateson claims in *Materials* that this type of variation was the primary force involved in speciation. (*Materials for the Study of Variation* 243)

that Pearson would be able to obtain the results he is trying to obtain using the methods he is employing. Second, even if such a value as a universal correlation constant were calculated, it would not be a meaningful descriptor of the processes of heredity.

In presenting his case against Pearson, Bateson begins with a step-by-step summation of the particular claims that the biometrician is making. He starts by restating Pearson's arguments concerning the assessment of homotypicity in organisms recalling the basic premises set forward by Pearson: 1) that homotypes, by definition, are a series of organs whose variability has its source in many small random influences and 2) that the homotypic correlation of undifferentiated like organs in a series can be lowered if variation in an organism has its source in differentiating influences (i.e. that two organs in a series have not been produced by the same process).

After laying out the mathematician's claims, Bateson challenges the biometrician's distinctions between *differentiation*, the occurrence of significant variation between two organs which leads to their being distinguished as two different types, and *variation*, the occurrence of slight changes within a homogeneous population of types which account for their distribution according to the law of error. He argues that from the outset Pearson's argument is in jeopardy because no clear distinction between these phenomena can be made in nature.

It is not, however, the difficulty of recognition I would now emphasize, but the fact that between the two phenomena [differentiation and variation] no absolute distinction exists in nature. An "undifferentiated series of like parts" means only a series of like parts which have varied and are varying among themselves but little.

A series of highly variable like parts is a series in which differentiation exists or is beginning to exist in a complex and irregular fashion. A “differentiated series of like parts” means a series among which variation is or has become definite of regular. Between these classes there is every shade and degree. No one can say finally where each begins and ends, and, by appropriate selection, we could find homotypic coefficients of any required value. The *average* value of such coefficients taken at random has no significance in nature. (“Heredity, Differentiation” 197)

In this section of the critique, Bateson’s criticism of Pearson’s biological model is a direct result of his own belief in Galton’s and Mendel’s models of inheritance, which hold that variation is the result of the random selection from a pool traits within the organism. Based on these models, different hereditary conditions can exist in different organisms at any given time in nature. In some populations of organs, there will be very little variation, because the population has reached what Galton would refer to as a place of stability for that organ. In other cases, however, there is shifting back and forth within the population between different places of stability. Finally, in still others two clearly distinct places of stability have emerged. The result is that across all organs there will be a wide rang of possible conditions with no clearly demarcated boundaries or values.

Because there is no way of knowing where variation ends and differentiation begins, Bateson argues that a researcher can select any set of values for correlation that they desire. As a result, they can come up with any average value for correlation they want. This, he argues, is what Pearson has done. He has simply selected a subset of

correlation values which yield the closest value to fraternal correlation. As a consequence, his contention that the correlation of undifferentiated like organs in nature is between .4 and .5 has no real purchase in nature.

Once he has made the case that the nature of heredity cannot be ascertained by statistical analysis because it is impossible to define the sample, Bateson makes the case that the only other alternative is to turn to close scrutiny of the individual. This position is, of course, in accordance with Mendel's and Bateson's approach to the study of heredity. It begins with gathering specific qualitative knowledge of individual instances through experimentation and then collecting and organizing the results of these experiments statistically to create a general picture of the hereditary process.

But let us now suppose that we could define differentiation from variation in general, say, as orderly variation. Even so we could not distinguish it unless its order was conspicuous.... Does not, then, the presence of orderly differentiation, in various degrees, *compel* us to an analysis of individual instances? In plain language, we shall have to pick and choose our cases, and the value of our coefficient of homotyposis will depend entirely on how we do it. (202)

For Bateson, the only way the appropriate sample can be found is through the development of qualitative knowledge of the organisms being observed. As an example of how this biologically-oriented experimental technique might be applied to the study of these phenomena, Bateson reconfigures Pearson's statistical observations of leaves on trees in terms of a qualitative experimental observation.

He [Pearson] speaks of the extreme difficulty of determining whether his material is homogenous in respect of the environment, but I miss from his work any deep appreciation of the subtle and evasive quality of differentiation. If anyone would obtain a conception of this difficulty, let him go to any tree or large plant and set about pruning it, or better, let him try to choose the shoots for propagation. Until he tries, it seems simple enough; but when he begins, he finds the shoots are of many complexly differing kinds, and unless he has experience pruning or propagating, he will not know which to choose. If he studies the tree attentively, he will soon see that the kinds of shoots are largely definite and, in fact, differentiated. (“Heredity, Differentiation” 202)

Bateson suggests here that if Pearson simply took time to study the individual instances in detail through experimental trial and error, as any naturalist worth his salt would do, he would soon be able to recognize which shoots represented distinct types and which did not. The results of this experimentation would lead him inevitably to Bateson’s conclusion that different kinds don’t develop in a series of small connected changes but rather from radical character saltations. From there, Pearson might then be able to construct a viable mathematical theory of heredity.

Such a theory, Bateson suggests, would not be statistical in the grand sense that it would seek, as Pearson’s theory does, to statistically gather all variation big and small together to find a mean value for the degree of variation. Instead, it would involve identifying points of stability, the degree of variation around those points of stability, and

places of discontinuity. This approach would describe variation and explore its range not to distill it into a meaningless universal average value.

By the one word *variation* we are attempting to express a great diversity of phenomena in their essence distinct though merging insensibly with each other. The attempt to treat or study them as similar is leading to utter confusion in the study of evolution.

If normality thus imagined can be shown a real phenomenon it is conceivable that we might then profitably attempt to determine in specified cases the average value of homotypic correlation for each case, but the average value for a miscellaneous collection of cases would still have no natural significance.

(203-04)

In Bateson's critique of Pearson's paper on homotyposis, it is possible to see both the power and the limits of mathematics in biological arguments. Just as Darwin, Mendel, and Galton recognized the power and the precision of quantification and mathematical formulae, operations, and principles, Bateson too recognizes their strength and does not attempt to make an issue of Pearson's application of correlation and the law of error to the study of heredity. What he does take issue with in his arguments, however, are the inconsistencies that he believes exist between Pearson's mathematical model and the actualities of nature. In taking this rhetorical approach, he argues not against Pearson's use of mathematics per se, but rather against Pearson's reduction of variation to a single average value, warranted by the law of large numbers. His opposition is predicated on his

belief in the Galtonian and Mendelian models of heredity in which traits can be in different modes of stability and in which sudden radical changes in traits can take place making the concept of an average value of variation an unhelpful if not meaningless measure for describing variation.

As history reveals, Pearson's arguments for a homotypic theory of heredity ultimately failed to persuade biologists. This failure, however, was not the result of his employing the principles, operations, and formulae from probability and statistics in making his arguments; rather, it was the result of a general movement away from Darwinian continuous variation and toward Mendel's apparently discontinuous theory of heredity. That the mathematical argumentation was not responsible for his failure in the minds of other biologists can be evidenced by quoting again George Shull's evaluation of Pearson's arguments:

The development of mathematical methods by Professor Karl Pearson and a few others promised for a time the key to the riddles of evolution. This movement, led by a pure mathematician, developed a series of beautiful methods for the mathematical analysis of data and the comparison of variations. These methods are of the greatest importance when rightly used but owing to the almost invariable lack of an equally keen *biological* analysis, the applications of these methods have led to a largely spurious product. (Shull 226)

Conclusion

From his early work with physical mathematics and graphical calculation to his later development of a theory of heredity based on the mathematics of the error curve, Pearson consistently advocated the importance of a theoretical understanding of mathematics to the exploration of both physical and organic phenomena. This belief in the primacy of mathematics as a source for discovering knowledge about nature informs the developments of his arguments in “On the Principle of Homotyposis and its Relation to Heredity” in which he makes the case for a mathematical model of heredity based on the error curve supporting Darwin’s theory of evolution by continuous variation.

Though Pearson’s belief in the importance of mathematics was generally accepted by scientists of all stripes from physics to biology, despite Pearson’s complaints to the contrary, his zeal for theoretical mathematics as a source for inventing arguments about nature met serious resistance from both physical and biological researchers. In the case of his biological arguments, William Bateson challenged Pearson’s insistence on the mathematical average of correlation as a fundamental value for the process of heredity. Bateson’s complaint was not over whether lines of arguments developed from the principles of probability and statistics could be used to describe biological phenomena, but rather that Pearson’s mathematical model and its goals were based on a biological model that was not indicative of actual biological conditions.

In the end, Pearson’s theory was rejected and forgotten as Darwin’s theory of discontinuous inheritance was swept away in the early twentieth century in a tide of evidence supporting Mendelian heredity. The mathematical tools that Pearson helped to

develop for determining correlation and testing goodness of fit, however, endured. The survival of these mathematical applications in biology testify not only to the importance that mathematics had assumed by the last decade of the nineteenth century in making biological arguments, but also to the importance of Karl Pearson as an advocate for the value of statistics and probability as sources of invention and support for biological arguments. Without the heightened awareness of the importance of mathematics to biology and without the environment of competition which demanded a mathematical response to discontinuous theories of evolution, the mathematically described laws of the monk from Brünn may not have had the opportunity to shine. Though substantively Mendel's work had appeared limited in scope to hybridists and botanists for thirty five years because they were a detailed analysis of a few traits in a few species, with the dawn of the twentieth century and the rise of statistical biology, the hidden value of Pearson's broad, mathematically informed methods and laws began to be recognized and to transform biology.

Epilogue

The goal of this project has been to explore the rhetorically unexplored territory of the role of mathematics in making arguments in evolutionary biology from the middle of the nineteenth to the beginning of the twentieth century and to extend the analysis of the arguments of nineteenth century evolutionary biologists past Darwin to the works of Mendel, Galton, and Pearson. The analyses undertaken to meet these goals reveal some important trends in argumentation in evolutionary biology in the latter half of the nineteenth century which have been previously overlooked.

First, rhetorical analyses of the texts and contexts of the arguments of Darwin, Mendel, Galton, and Pearson suggest that they believed that their arguments could be strengthened with the deductive *logos* and *ethos* of mathematics. In Darwin's work, for example, he employed quantitative comparison as a persuasive tool for making the case for variation, natural selection, and the principle of divergence of character. Additionally, Mendel used combinatorics to develop experiments to prove the invariability of character traits over many generations of breeding.

A second trend revealed by these analyses is the importance of mathematical concepts as a source for inventing arguments about nature. In chapters four and five, for example, analyses of the arguments of Galton and Pearson illuminate the importance of the error curve and the principles of the law of error to develop new lines of arguments about the nature of variation, evolution, and heredity. Chapter four reveals how, by replacing the concept of "error" with "variation," Galton transformed the error curve into an instrument for describing the distribution of variation within human populations for any measurable physiological or psychological feature. Additionally, it illustrates how

this transformation opened up for him new paths of empirical investigation and encouraged the development of new methods for mathematically establishing the degree of change in quantifiable heritable features from one generation to the next which served as the foundational arguments for this law of ancestral heredity.

Similarly, chapter five shows how Pearson, inspired by Galton's arguments from the error curve, developed arguments for a phenomenological theory of heredity based on the law of error in which Galton's concept of "error as variation" was transformed back to Quetelet's original notion of "variation as error." This transformation brought arguments about the biological process of variation, error, and evolution completely under the jurisdiction of mathematical evidence and principles.

In addition to revealing important general trends in the role of mathematics in making arguments in evolutionary biology, this investigation also illuminates specific instances of the limitations of mathematical arguments in persuasion for some audiences. These limitations are brought to the surface in the analyses of the failure of Mendel's brilliantly constructed mathematical arguments in "Experiments in Plant Hybridization" and in Karl Pearson's mathematically sound but biologically ill-conceived theory of homotyposis.

The rhetorical analysis of Mendel's "Experiments" in chapter three reveals that though Mendel employs in his argument a carefully implemented system of mathematically informed experiments, his results were ignored by biological researchers in Western Europe. By analyzing the rhetorical context in which Mendel makes his arguments, the case was made that Mendel's conclusions, despite their underlying mathematical rigor, were ignored because he failed to take account of in his work, and in

fact argued against, biological phenomena which were widely accepted by biologists at the time. Without the proper professional ethos to attract supporters to his unorthodox position, Mendel's conclusions languished in obscurity for more than thirty years despite the fact that they were mathematically supported. Similar persuasive limitations were revealed by the rhetorical analysis of the context and content of Pearson's arguments in chapter five. This investigation concludes that his mathematical arguments also fail because they support a vision of heredity at odds with the biological evidence accepted by the majority of biologists at the time he is writing.

While my analyses of the content and context of Mendel and Pearson's arguments provide evidence that there are persuasive limitations for mathematical arguments, my investigation of Galton's arguments explores how other rhetorical circumstances and strategies can furnish the necessary support to persuade readers to accept potentially controversial mathematical arguments. My analysis in chapter four of Galton's position in the social and scientific community suggests that the success of his arguments for an analogy between the law of error, variation, and heredity can be attributed in part to the strength of his ethos and in part to his direct persuasive efforts. The investigation of Galton's personal and professional history in the chapter reveals that he was a well regarded member of the highest social and intellectual circle and, therefore, well positioned to inspire in British scientific audiences a willingness to consider his arguments in *Natural Inheritance*. In addition, a rhetorical analysis of his arguments in *Natural Inheritance* provides evidence that some of Galton's success can be attributed to his skill as a rhetor. This skill is evident in the second and third chapters in which he deftly argues for the jurisdiction of the error curve in arguments about variation, heredity,

and evolution. Galton's powers of persuasiveness are also illustrated in the later chapters of his text where he appeals to the prevailing social and scientific sentiment of the audience to persuade them to accept the analogy between the error curve and variation in nature.

Though this project has focused specifically on the role of mathematics in making arguments in evolutionary biology and the success or failure of these arguments for specific historical figures in particular rhetorical contexts, the use of mathematics in biology, in particular the error curve, has important implications beyond the current limits of this investigation. Future areas of rhetorical exploration might include the role of the error curve in the debates over fate and free will in the middle of the nineteenth century, the influence of the error curve in the development of arguments for eugenics in the end of the nineteenth century, and the use of the error curve as support for communist ideology in the latter half of the nineteenth century.

Although the debate over whether fate or free will governed human life has been a staple topic of disagreement at least as far back as Plato, at different times in history this debate has dominated the public forum. In the middle of the nineteenth century, an increased interest in the debate developed following the publication of H.T. Buckle's *History of Civilization in England* in which Buckle used the error curve to support a historical argument that the course of human societies was strictly determined by climate and locality (Chaterjee 278).

A rhetorical examination of Buckle's work would offer further detail about the social and rhetorical dimensions of the error curve in the nineteenth-century public imagination in England. Importantly, it might give further clues about the involvement of

the mathematical construct in the complex debates over causation which developed with the emergence of positivist philosophy in the late nineteenth and early twentieth centuries.

In addition to controversies over personal fate and free will, the error curve as a rhetorical warrant also plays a role in debates about the degree to which societies can control their evolutionary destiny. Fears of eradication resulting from evolutionary change seem to be the central impetus around the formation of the eugenics movement in England in the latter part of the nineteenth century. Both Francis Galton and Karl Pearson play seminal roles in articulating and promoting the precepts of eugenics. Given their interest in the error curve and their belief in its power to describe the limit and distribution of variation in organic populations, it is more than likely that the curve played a role in developing and supporting eugenic arguments.

The importance of eugenics to social and political debates of the late nineteenth and early twentieth century encourages a more detailed rhetorical examination of the nature and degree of influence of the error curve in developing eugenic arguments. Such an investigation might profitably explore how the concept of “true value” might have affected a new definition for a biological ideal. It might also investigate how the concept of regression and the possibility of a shifting population mean may have provided inspiration for the necessity of and possibility for a eugenic program of social improvement.

In opposition to some elements of Galtonian eugenic sentiment, the error curve may have also been a source for arguments supporting communist ideology. One of the major lines of argument that communist ideologues sought to discredit was that the

structure of the social class was simply the expression of an innate biological superiority of some blood lines. The appearance of the law of error, which allowed for the mathematical description and comparison of biological traits, provided the opportunity to challenge, if not to refute, elitist arguments for biological determinism. An investigation of the role of the error curve and concepts such as regression spawned from the curve might offer some useful insights into the role of biology and bio-mathematical arguments in communist rhetoric. It might also open a useful line of investigation into the relationship between the eugenics and the communist movements, particularly in the arguments of Karl Pearson, who was an ardent supporter of both (Porter 2004).

The number and variety of these lines of investigation underscore the importance of the error curve as a guiding topic for developing arguments in the latter half of the nineteenth century. In addition, they reveal the importance of mathematics in the nineteenth century as a source for arguments not only in science, but also as an influence on some of the most important socio-political movements and debates of the period. Although the role of mathematics in argumentation currently receives slight attention in rhetorical scholarship, the possibilities for rhetorical investigation offered by the error curve in the latter half of the nineteenth century alone should provide encouragement that a broad and virtually untapped source for rhetorical investigation is waiting to be explored.

Appendix A

List of Scholarly Books Analyzed for Citations of the Role of Mathematics in Darwin's Arguments.

- 1) Gale, Berry G. *Evolution without Evidence Charles Darwin and the Origin of Species*. Albuquerque: U of New Mexico P, 1982.
 - 2) Ghislen, Michael T. *The Triumph of the Darwinian Method*. Berkeley: U of California P, 1969.
 - 3) Gillespie, Neal C. *Charles Darwin and the Problem of Creation*. Chicago: U of Chicago P, 1979.
 - 4) Brent, Peter. *Charles Darwin a Man of Enlarged Curiosity*. New York: Harper and Row, 1981.
 - 5) Bowlby, John. *Charles Darwin a New Life*. New York: Norton, 1991.
 - 6) Browne, Janet. *Charles Darwin the Power of Place*. Vol. 2. Princeton: Princeton UP, 1996.
 - 7) Bowler, Peter. *Charles Darwin the Man and his Influence*. Cambridge: Cambridge UP, 1990.
 - 8) Hull, David. *Darwin and His Critics: The Reception of Darwin's Theory of Evolution by the Scientific Community*. Cambridge: Harvard UP, 1973.
 - 9) Aydon Cyril. *Charles Darwin: The Naturalist Who Started a Scientific Revolution*. New York: Carroll and Graf Publishers, 2002.
 - 10) Mayr, Ernst. *Charles Darwin and the Genesis of Modern Evolutionary Thought*. Cambridge: Harvard UP, 1991.
 - 11) Campbell, John Angus. "Charles Darwin Rhetorician of Science." *The Rhetoric of the Human Sciences Language and Argument in Scholarship and Public Affairs*. Ed. John S. Nelson. Madison: The U of Wisconsin P, 1987. 69-86.
-
- 1) "Mathematics" in the index: p. 13-14 discusses Darwin's weakness in mathematics. Page 79 explains that Darwin read *Nature* but that he did not understand most of the articles which used mathematics. / Gale discusses the contents of Darwin's arguments in depth. He looks at his arguments, their weaknesses and the problems of evidence in chapter 7 of the text. At no point does he discuss the role of mathematics in presenting Darwin's evidence.
 - 2) "Mathematics" in the index: p21 Defends Darwin's not using mathematics to make his geological arguments. He argues that something can be logical whether it is expressed in a language of formal logic like mathematics or not. P65 Again he attacks mathematical argument in the harder sciences. P127 Darwin uses geometry to describe changes in the physical shape of organisms. P170 Darwin quantifies different physical features in organisms to present precise morphological comparisons. P175 Discusses Darwin's hybridization experiments. It claims that his results were subjected to statistical analysis by Galton (though its not clear whether this was done before or after publication)
 - 3) Gillespie doesn't mention mathematics in the index. He doesn't discuss it with regards to Darwin's approach to his subject.

Appendix A (continued)

- 4) Brent doesn't mention mathematics in the index of his text. In skimming chapters 4 and 5 "Origin of all My Views" and "The Speculatist" I found no discussion of mathematics.
- 5) Bowlby doesn't mention mathematics in the index of his text. Chapter 22 "My Abominable Volume" discusses generally the contents of the argument in Darwin's text. It does not, however, go into specifics. Also, it does not include any mention of mathematics.
- 6) Browne doesn't mention mathematics in the index of her text. She talks at length in chapter 2 "My Abominable Volume" about Darwin's use of evidence but at no time discusses the role of mathematic notation in presenting that evidence.
- 7) Bowler doesn't mention mathematics in the index of his text. On P104 of his text he mentions that Darwin undertakes arithmetical studies of the number of species to prove that population pressure within a given area forces species to diversify in order to compete more successfully. On P 117 he also discusses the importance of mathematical reasoning in Malthus' argument which informs his concept of natural selection.
- 8) Hull has an entry for mathematics with 6 page references in the index. On P 11 he comments how Darwin was no more confused about the differences between scientific laws, mathematical axioms, and metaphysical principles than most people in the period in which he lived. Hull claims that despite his misunderstanding of these differences Darwin understood the general distinction between deductive and inductive reasoning. On P 14 Hull explains that Darwin was caught in the middle of a debate about the difference between, "the nature of mathematical axioms and their relation to their experience". On page 20 Hull argues that mathematics for Aristotle was different from empirical science because it required fewer observations to reach a conclusion. On p 32 Hull claims that Darwin's theory is not mathematical in form or in the type of reasoning. P34 Hull claims that in the 1930s different scientist tried through genetics to create a mathematical science of evolutionary biology. This was not effective in creating verifiable outcomes, however. P61 Hull argues that evolution did not exhibit mathematical regularity as did the laws in hard science.
- 9) Aydon doesn't mention mathematics in the index of his text. His only reference to it is in his comment that Darwin was not particularly good in mathematics in school on P 30.
- 10) Mayr doesn't mention mathematics in the index of his text. In chapter 5 p48 Mayr notes that the concept of science in Darwin's time was completely dominated by mathematics and the physical sciences.
- 11) Campbell does not discuss the role of mathematics in his chapter of the text.

Appendix B

List of Scholarly Journals Analyzed for Citations of the Role of Mathematics in Darwin's Arguments.

- 1) Browne, Janet. "Darwin's Botanical Arithmetic and the 'Principle of Divergence,' 1854-1858". *Journal of the History of Biology* 13 (1980): 53-89.
- 2) Campbell, John Angus. "Why was Darwin Believed? Darwin's *Origin* and the problem of Intellectual Revolution." *Configurations* 11.2 (2003): 203-37.
- 3) Campbell, John Angus. "The Invisible Rhetorician: Charles Darwin's 'Third Party' Strategy." *Rhetorica* 7.1 (1989): 55-85
- 4) Moore, Randy. "The Persuasive Mr. Darwin." *Bioscience* 47.2 (1997): 107-14.
- 5) Gould, Stephen J. "False Premise Good Science." *Natural History* 92.10 (1983): 20-26.
- 6) Manier, Edward. "'External Factors' and 'Ideology' in the Earliest Drafts of Darwin's Theory." *Social Studies of Science* 17.4 (1987): 581-609.

- 1) Complete discussion of the role that mathematics in the form of botanical arithmetic played in developing Darwin's principle of divergence in species.
- 2) No discussion of mathematics. Campbell is more interested in Darwin's linguistic choices than in his use of evidence to support his arguments.
- 3) Does not discuss mathematics.
- 4) Moore recognizes the role of data in making Darwin's arguments but not the role of mathematical operations or principles (pp107-108). He focuses primarily on the linguistic strategies that Darwin uses to present his arguments.
- 5) Gould discusses the debate about the age of the earth between Darwin and Kelvin. His point is that calculations aren't everything and instead it's the frame of thinking. There is little discussion of Darwin's argument except to say that he stuck to his guns for a long time regarding the age of the earth. It doesn't provide any in-depth discussion of Darwin's argument per se.
- 6) On page 589 Manier suggests that Darwin rejected, "The Cartesian criteria for mathematical clarity and distinctness in biological concepts". Here he suggests as Hull does that Darwin's arguments are in form non-mathematical. This, however, does not preclude the use of mathematics in his argument. He also suggests (P.529) that Darwin did not rely on empirical evidence. There is a two-fold definition of empirical evidence in Webster's which is something, "capable of being verified or disproved by observation or experiment". I believe that Darwin's observations were verifiable in the sense that he had testimonials from others who said that what he witnessed was so. Later Manier argues that Darwin was a pioneer in statistical thinking and that his arguments were based on statistical probability rather than *vera causa*. This comment focuses on the mode of argument rather than the means of presenting evidence.

Appendix C

Mathematical Tables Created by Darwin Used in *The Variation of Animals and Plants under Domestication*

Charles Darwin, *The Variation of Animals and Plants under Domestication*, 2 vols. 2nd ed. (New York: D. Appleton & Co., 1883) 175, 182.

Page 175

The following table will serve as a summary, and will show the most remarkable deviations in the number of the vertebra and ribs which I have observed:—

	Rock Pigeon.	Pouter, from Mr. Bult.	Tumbler, Dutch Roller.	Bussorah Carrier.
Cervical Vertebrae	12	12	12	12 The 12th bore a small rib.
Dorsal Vertebrae	8	8	8	8
" Ribs ..	8 The 6th pair with processes, the 7th pair without a process.	8 The 6th and 7th pair with processes.	7 The 6th and 7th pair without processes.	7 The 6th and 7th pair without processes.
Sacral Vertebrae	12	14	11	11
Caudal Vertebrae	7	8 or 9	7	7
Total Vertebrae	39	42 or 43	38	38

Appendix C (continued)

Page 182

TABLE II.

Pigeons with their beaks longer than that of the Rock-pigeon, proportionally to the size of their bodies.

Name of Breed.	Actual length of Feet	Difference between actual and calculated length of feet, in proportion to length of feet and size of body in the Rock-pigeon.	
		Too short by	Too long by
Wild rock-pigeon (mean measurement)	2.02	Too short by	Too long by
Short-faced Tumbler, bald-head	1.57	0.11	..
Carrier	2.60	..	0.31
Carrier	2.60	..	0.25
Carrier	2.40	..	0.21
Carrier Dragon	2.25	..	0.06
Bagadotten Carrier	2.80	..	0.56
Scanderoon, white	2.80	..	0.37
Scanderoon, Pigeon cygne	2.85	..	0.29
Runt	2.75	..	0.27
Number of specimens	8	..	8

Appendix D

A Mathematical Table Created by Darwin Used in *Natural Selection*

Charles Darwin, *Charles Darwin's Natural Selection*, ed. R.C. Stauffer (Cambridge: Cambridge UP, 1975) 149.

VARIATION UNDER NATURE			
TABLE I ¹			
<p>For particulars on the works here tabulated and on the few corrections made, see the Supplement to this Chapter.</p>	<p>The numerators in the columns give the number of species presenting varieties; the denominators the number of species in the larger and smaller genera: these fractions are all reduced to common denominators of a thousand for comparison, and are printed in larger type to catch the eye. The right hand rows of figures in the three columns, with decimals, show the average number of varieties which each varying species has,— thus the number 1.50 shows that each two varying species have on average between them three varieties.</p>		
	<i>Larger Genera</i>	<i>Smaller Genera</i> (including those with single species)	<i>Genera with a single species</i>
Great Britain. Bentham			
Great Britain: Babington —Larger Genera with 5 species and upwards, smaller with 4 species and downwards [Pencil note by C.D.: Write this column larger'.]	$\frac{101}{663} = \frac{152}{1000} \quad 1.40$	$\frac{89}{745} = \frac{119}{1000} \quad 1.30$ [Pencil note by C.D.: 'Write this larger'.]	$\frac{24}{255} = \frac{94}{1000} \quad 1.50$
Great Britain, Henslow— Larger Genera with 5 species and upwards, smaller with 4 species and downwards. The Varieties are divided into two groups, the less strongly marked, and those which have been ranked by some eminent Botanists as species. Lesser Vars:	$\frac{69}{560} = \frac{123}{1000} \quad 1.55$	$\frac{67}{692} = \frac{96}{1000} \quad 1.40$	
Stronger Vars:	$\frac{33}{560} = \frac{58}{1000} \quad 1.33$	$\frac{29}{692} = \frac{41}{1000} \quad 1.20$	
Great Britain—London Catalogue (1853) (see Supplement for nature of Varieties)—Larger Genera with 5 species and upwards, smaller with 4 species and downwards	$\frac{97}{616} = \frac{157}{1000} \quad 1.35$	$\frac{85}{642} = \frac{132}{1000} \quad 1.27$	
Great Britain—London Catalogue—forms ranked as species in this catalogue but which have been thought by some authors to be varieties. In this second line, larger genera with 5 species and upwards, smaller with 4, 3, and 2 species	$\frac{57}{559} = \frac{101}{1000}$	$\frac{14}{377} = \frac{37}{1000}$	

¹ [Darwin's holograph draft for this table is in ULC vol. 16.1, fol. 167.]

Appendix E

Mathematical Tables Created by Darwin Used in *Natural Selection*

Charles Darwin, *Charles Darwin's Natural Selection* ed. R.C. Stauffer (Cambridge: Cambridge UP, 1975)150-51.

Page 150

VARIATION UNDER NATURE

Table 1 cont.

	<i>Larger Genera</i>		<i>Smaller Genera</i> (including those with single species)		<i>Genera with a single species</i>
Centre France: Boreau— Larger Genera with 5 species and upwards, smaller with 4 species and downwards.	$\frac{113}{732} = \frac{154}{1000}$	1.38	$\frac{84}{741} = \frac{107}{1000}$	1.47	$\frac{19}{267} = \frac{721}{1000}$ 1.47
Holland: Miquel—Larger Genera with 4 species and upwards, smaller with 3 species and downwards.	$\frac{22}{622} = \frac{35}{1000}$		$\frac{25}{557} = \frac{44}{1000}$		
Germany & Switzerland: Koch—Larger Genera with 7 species, and upwards, smaller with 6 species and downwards	$\frac{390}{2093} = \frac{186}{1000}$	1.72	$\frac{162}{1365} = \frac{118}{1000}$	1.79	$\frac{32}{345} = \frac{92}{1000}$ 1.50
Dalmatia: Visiani—Larger Genera with 5 species and upwards, smaller with 4 species and downwards.	$\frac{164}{1007} = \frac{162}{1000}$	1.37	$\frac{130}{899} = \frac{144}{1000}$	1.31	$\frac{46}{290} = \frac{158}{1000}$ 1.26
Rumelia: Grisebach— Larger Genera with 6 species and upwards, smaller with 5 species and downwards.	$\frac{98}{1136} = \frac{86}{1000}$	1.45	$\frac{54}{1083} = \frac{49}{1000}$	1.14	$\frac{12}{326} = \frac{36}{1000}$ 1.16
Russia, Ledebour (All 4 vols together) Larger Genera with 10 species and up- wards, smaller with 9 species and downwards	$\frac{692}{3955} = \frac{174}{1000}$	1.48	$\frac{307}{2407} = \frac{127}{1000}$	1.39	$\frac{45}{475} = \frac{94}{1000}$ 1.26
Ledebour—Vol: I separately.	$\frac{207}{1237} = \frac{167}{1000}$	1.42	$\frac{62}{576} = \frac{107}{1000}$	1.32	
———— Vol: II ————	$\frac{192}{1243} = \frac{154}{1000}$	1.56	$\frac{94}{767} = \frac{122}{1000}$	1.35	
———— Vol: III ————	$\frac{171}{905} = \frac{188}{1000}$	1.49	$\frac{94}{595} = \frac{157}{1000}$	1.50	
———— Vol: IV ————	$\frac{122}{570} = \frac{214}{1000}$	1.45	$\frac{57}{470} = \frac{121}{1000}$	1.36	
N. United States. A. Gray— Larger Genera with 5 species and upwards, smaller with 4 sp. and downwards. The two kinds of vars. marked in this work are here classed together.	$\frac{112}{1136} = \frac{98}{1000}$	1.40	$\frac{65}{917} = \frac{70}{1000}$	1.36	$\frac{32}{361} = \frac{88}{1000}$ 1.37

Appendix E (continued)

Page 151

VARIATION UNDER NATURE

Table 1 cont.

	<i>Larger Genera</i>		<i>Smaller Genera</i> (including those with single species)		<i>Genera with a single species</i>
Canary Islands, Webb & Berthelot—Larger Genera with 4 species and upwards, smaller with 3 and downwards.	$\frac{49}{421} = \frac{116}{1000}$		$\frac{42}{551} = \frac{76}{1000}$		
India (part of Flora) Hooker & Thomson—Larger Genera with 7 species and upwards, smaller with 6 species and downwards.	$\frac{21}{258} = \frac{81}{1000}$ 1.61		$\frac{13}{165} = \frac{78}{1000}$ 1.53		
Tierra del Fuego: Hooker—Larger Genera with 3 species and upwards, smaller with 2 species and downwards	$\frac{19}{177} = \frac{107}{1000}$ 1.57		$\frac{16}{163} = \frac{98}{1000}$ 1.37		
New Zealand: Hooker—Larger Genera with 4 species and upwards, smaller with 3 species and downwards	$\frac{52}{361} = \frac{149}{1000}$ 1.82		$\frac{37}{323} = \frac{114}{1000}$ 2.05		$\frac{15}{159} = \frac{94}{1000}$ 2.00
Insecta: Coleoptera Madeira: Wollaston—Larger Genera with 4 species and upwards, smaller with 3 species and downwards	$\frac{35}{225} = \frac{155}{1000}$ 1.71		$\frac{26}{257} = \frac{101}{1000}$ 1.34		
Sweden-Gyllenhal—Larger Genera with 11 species and upwards, smaller with 10 species and downwards	$\frac{512}{1344} = \frac{380}{1000}$ 1.85		$\frac{151}{485} = \frac{311}{1000}$ 1.43		$\frac{11}{43} = \frac{255}{1000}$ 1.54

that there were many great difficulties in the way. The subject is so highly important to us, as we shall see in a future chapter, that these difficulties must be discussed at tedious length; but it will be convenient first to give the tables./

A 26/In Table 1, we have several of the best known local Floras, (some of which were selected for me by Dr. Hooker) with the species divided into two great groups, those in the larger & those in the smaller genera. On the extreme right hand we have the genera with only a single species, but these are likewise included amongst the smaller genera. Some of the smaller Floras have been selected simply from giving remote countries under different climates. I may premise that I have given every single Flora (&

Appendix F

An Example of Kölreuter's Graphical Presentation of Crosses.

Joseph Gottlieb Kölreuter, *Vorläufige Nachricht von einigen das Geschlecht der Pflanzen betreffenden Versuchen und Beobachtungen*, Microform (Leipzig: Gleditschischen Handlung, 1761-1766. Volume 4) 48.

XXII. Vers.
 Nicot. rust. ♀.
 rust. ♀. }
 panic. ♂. } ♀.
 Nicot. }
 panic. ♂. }
 panic. ♂. } ♂.

Von diesem Versuche wurden sechs Pflanzen erzogen. Ich konnte zwischen ihnen und den einfachen aus der rust. ♀ und panic. ♂ erzeugten Bastarten keinen merklichen Unterschied finden.

XXIII. Vers.
 Nicot. rust. ♀. }
 panic. ♂. } ♀.
 rust. ♀. }
 panic. ♂. } ♀.
 Nicot. }
 panic. ♂. }
 panic. ♂. } ♂.

Ich erzog hievon drey Pflanzen. Eine derselben war ihrer ganzen äußerlichen Anlage nach dem in der zweyt, Forts. S. 16. S. 73. 2c. beschriebenen Bastart im ersten aufsteigenden Grade sehr ähnlich, und hinterließ viele, aber ganz leere Kapseln. Die zwey übrigen hatten etwas weniger Ähnlichkeit mit der panic. als die erstern, und setzten nur sehr wenig, ziemlich spitziae und ebenfalls ganz leere Kapseln an. Man sieht hieraus

Appendix G

An Example of Quantitative Comparison in Kölreuter's Work

Joseph Gottlieb Kölreuter, *Vorläufige Nachricht von einigen das Geschlecht der Pflanzen betreffenden Versuchen und Beobachtungen*, Microform (Leipzig: Gleditschischen Handlung, 1761-1766. Volume 1) 41.

Zur Seite 41. Vergleichungsmaß.

	Nicot. rust.	rust. ♀ panic. ♂	panic.	rust. ♀ α. panic. ♂ peren. ♂	rust. ♀ β. pan peren
Länge der ganzen Blume von dem Grunde der Blumenröhre an bis zu dem nach aufwärts getretenen und in fünf Einschnitte abgetheilten Blumenrande:	7'''	9 ² / ₇ '''	1'', 1 ¹ / ₂ '''	1'', 3'''	1''
Länge des Blumenfelds: von seinem Grunde an bis an die Spitze des längsten Einschnitts:	5 ¹ / ₂ '''	5 ¹ / ₂ '''	3 ² / ₇ '''	7 ¹ / ₂ '''	7
Die Blume ragt über die Spitze des längsten Reifeinschnitts heraus:	1 ¹ / ₂ '''	4'''	9 ¹ / ₂ '''	6 ¹ / ₂ '''	1'',
Größte Breite von einem Ende des ganzen Blumenrandes bis zum andern, quer über die Blume gemessen:	7 ¹ / ₇ '''	5 ¹ / ₇ '''	4 ¹ / ₇ '''	9'''	10
Breite (oder vielmehr Länge) des absteigenden Blumenrandes selbst:	2 ⁵ / ₇ '''	2'''	1 ¹ / ₄ '''	3 ¹ / ₂ '''	5
Durchmesser der Blumenröhrenöffnung zwischen dem Rande:	2 ¹ / ₄ '''	2 ¹ / ₄ '''	1 ² / ₇ '''	3'''	3
Durchmesser des Blumenröhrenbauchs unter dem Rande:	3 ¹ / ₂ '''	3'''	2 ¹ / ₂ '''	4'''	3
Ganze Länge der Blumenröhre:	6 ² / ₇ '''	9'''	1'', 2 ¹ / ₇ '''	1'', 1 ¹ / ₂ '''	1''
Länge des engen Grundes der Blumenröhre:	1 ¹ / ₂ '''	2 ¹ / ₂ '''	3'''	4 ¹ / ₂ '''	6
Länge der Staubfäden:	4'''	5 ¹ / ₂ '''	8 ¹ / ₇ '''	9 ¹ / ₂ '''	1

Appendix H

Table of Massaged Measures from Natural Inheritance

Francis Galton, *Natural Inheritance* (London: Macmillan and Company, 1889) 201.

TABLE 3.

DEVIATIONS from \bar{M} in each of the series in Table 2, after reduction to a Scale in which $Q' = 1$, where Q' is the *Mean* of the observed Deviations at the Grades 20°, 30°, 70°, and 80°.

Subject of measurement.	Values of Q'	Unit of measurement in Table 2.	Sex.	No. of persons	Deviations reckoned in units of Q' .											
					5°	10°	20°	30°	40°	50°	60°	70°	80°	90°	95°	
Height, standing, without shoes . . .	1.72	Inches	{	M.	811	2.73	1.98	1.22	0.81	0.35	0	0.35	.76	1.22	1.98	2.61
	1.62			F.	770	2.71	2.10	1.23	.74	.37	0	.37	.80	1.23	1.91	2.46
Height, sitting, from seat of chair . . .	0.95	Inches	{	M.	1013	2.52	1.89	1.15	.73	.63	0	.31	.73	1.15	1.79	2.31
	0.82			F.	775	2.55	1.95	1.22	.73	.36	0	.36	.85	1.22	2.07	2.55
Span of arms . . .	2.07	Inches	{	M.	811	2.36	1.83	1.30	.82	.43	0	.33	.72	1.16	1.79	2.36
	1.87			F.	770	2.35	1.87	1.23	.69	.32	0	.37	.80	1.28	1.98	2.67
Weight in ordinary indoor clothes . . .	10.00	Pounds	{	M.	520	2.20	1.80	1.20	.80	.40	0	.40	.70	1.30	2.20	2.90
	11.00			F.	276	1.80	1.60	1.10	.70	.40	0	.60	.90	1.30	1.80	2.40
Breathing capacity . . .	24.50	Cubic	{	M.	212	2.32	1.68	1.28	.80	.32	0	.28	.68	1.16	2.32	2.84
	19.00			F.	277	2.39	1.87	1.20	.73	.36	0	.31	.67	1.35	2.03	2.49
Strength of pull as archer with bow . . .	7.50	Pounds	{	M.	519	2.39	1.86	1.33	.80	.40	0	.40	.80	1.06	1.99	2.92
	5.22			F.	276	1.92	1.06	.80	.53	.27	0	.27	.53	.93	1.46	1.86
Strength of squeeze with strongest hand	7.75	Pounds	{	M.	519	2.32	1.81	1.16	.77	.39	0	.39	.77	1.29	1.93	2.45
	7.50			F.	276	2.12	1.73	1.20	.66	.40	0	.40	.80	1.33	1.99	2.66
Swiftmess of blow	2.37	Ft. per second	{	M.	516	2.06	1.68	1.22	.80	.34	0	.42	.80	1.18	1.77	2.31
	1.55			F.	271	2.71	2.13	1.35	.84	.38	0	.38	.71	1.10	1.87	2.26
Sight, keenness of — by distance of reading diamond test-type	4.00	Inches	{	M.	398	3.00	2.00	1.25	.75	.50	0	.25	.75	1.25	1.75	2.25
	5.22			F.	433	2.66	2.28	1.52	.95	.38	0	.38	.57	.95	1.33	1.52
SUMS . . .					43.11	33.12	21.96	13.65	7.00	0	6.57	13.34	21.46	33.96	43.82	
MEANS . . .					2.40	1.84	1.22	0.76	0.39	0	0.37	0.74	1.19	1.89	2.43	
MEANS multiplied by 1.015, to change unit to $Q = 1$					2.44	1.87	1.24	0.77	0.40	0	0.38	0.75	1.21	1.92	2.47	
Normal Values, when $Q = 1$					2.44	1.90	1.25	0.78	0.38	0	0.38	0.78	1.25	1.90	2.44	

Appendix I

Galton's Table of the "Frequency of Different Eye Colors in Four Successive Generations"

Francis Galton, *Natural Inheritance* (London: Macmillan and Company, 1889) 212.

TABLE 15.
FREQUENCY OF DIFFERENT EYE-COLOURS IN FOUR SUCCESSIVE GENERATIONS.

Sex and the No. of the (ascending) generation.	No. of cases of eye-colour observed.								Percentages.										
	1. Light blue.	2. Blue. Dark blue.	3. Grey. Blue-green.	4. Dark grey. Hazel.	5. Light brown.	6. Brown.	7. Dark Brown.	8. Very dark brown. Black.	Totals.	1. Light blue.	2. Blue. Dark blue.	3. Grey. Blue-green.	4. Dark Grey. Hazel.	5. Light brown.	6. Brown.	7. Dark brown.	8. Very dark brown. Black.	Totals.	
Males {	IV.....	13	177	136	40	2	39	44	12	463	2.8	38.2	29.4	8.6	0.4	8.4	9.5	2.6	99.9
	III.....	19	234	233	84	3	79	97	24	773	2.4	30.3	30.1	10.9	0.4	10.1	12.6	3.1	99.9
	II.....	30	167	236	108	8	83	74	36	742	4.0	22.5	31.8	14.6	1.1	11.2	10.0	4.8	100.0
	I.....	3	89	82	47	1	37	31	9	299	1.0	28.9	27.4	15.7	0.3	12.4	10.4	3.0	100.0
General.....	65	687	687	279	14	238	246	81	2277	2.9	29.3	30.2	12.3	0.6	10.4	10.8	3.6	100.0	
Females {	IV...	7	132	114	48	2	70	58	19	450	1.5	29.3	25.3	10.7	0.4	15.6	12.9	4.2	99.9
	III..	22	173	241	89	7	100	96	17	742	2.9	23.3	32.5	12.1	0.9	13.5	12.5	2.3	100.0
	II..	21	210	241	98	3	78	60	24	735	2.9	28.6	32.8	13.3	0.4	10.6	8.2	3.3	100.1
	I.....	6	78	82	55	5	33	22	5	286	2.1	27.3	28.7	19.2	1.7	11.5	7.7	1.7	99.0
General.....	56	593	678	290	17	281	233	65	2213	2.5	26.8	30.6	13.1	0.8	12.7	10.5	2.9	99.9	
Males and Females {	IV...	20	309	240	88	4	109	102	31	913	2	34	27	10	1	12	11	3	100
	III..	41	407	474	173	10	179	190	41	1515	3	27	31	11	1	12	12	3	100
	II..	51	377	477	206	11	161	134	60	1477	3	26	32	14	1	11	9	4	100
	I.....	9	167	164	102	6	70	53	14	585	1	29	28	18	1	12	9	2	100
General.....	121	1260	1365	569	31	519	479	146	4490	2.7	28.1	30.4	12.7	0.7	11.6	10.7	3.3	100.2	

Appendix J

Galton's Tables Comparing the Observed and Calculated Frequency of Eye Color in 16 Family Groups

Francis Galton, *Natural Inheritance* (London: Macmillan and Company, 1889) 215.

TABLE 19.

OBSERVED AND CALCULATED EYE-COLOURS IN 16 GROUPS OF FAMILIES.

Those families are grouped together in whom the distribution of Light, Hazel, and Dark Eye-colour among the Parents and Grandparents is alike. Each group contains at least Twenty Brothers or Sisters.

Eye-colours of the						Total child- ren.	Number of the light eye- coloured children.			
Parents.			Grandparents.				Ob- served.	Calculated.		
Light.	Hazel.	Dark.	Light.	Hazel.	Dark.			I.	II.	III.
2	4	183	174	161	163	172
2	3	1	...	53	46	47	44	48
2	3	...	1	92	88	81	67	79
2	2	1	1	27	26	24	18	22
...	...	2	2	...	2	22	11	6	12	6
1	1	...	3	1	...	62	52	48	51	51
1	1	...	3	...	1	42	30	33	31	32
1	1	...	2	2	...	31	28	24	24	20
1	1	...	2	...	2	49	35	38	28	34
1	1	...	2	1	1	31	25	24	21	23
1	...	1	3	...	1	76	45	44	55	46
1	...	1	2	...	2	66	30	38	38	35
1	...	1	2	...	1	27	15	16	18	16
1	...	1	1	...	3	20	9	12	8	9
1	...	1	1	1	2	22	8	13	11	11
...	1	1	1	1	2	24	9	14	12	10
							629	623	601	614

Appendix K

Galton's Tables Comparing the Observed and Calculated Frequency of Eye Color in Several Generations

Francis Galton, *Natural Inheritance* (London: Macmillan and Company, 1889) 216-17.

Page 216

TABLE 20.
OBSERVED AND CALCULATED EYE-COLOURS IN 78 SEPARATE FAMILIES, EACH
OF NOT LESS THAN SIX BROTHERS OR SISTERS.

Eye-colours of the						Total child- ren.	Number of the light eye- coloured children.			
Parents.			Grandparents.				Ob- served.	Calculated.		
Light.	Hazel.	Dark.	Light.	Hazel.	Dark.			I.	II.	III.
2	4	6	6	5.3	5.3	5.6
2	4	6	6	5.3	5.3	5.6
2	4	6	6	5.3	5.3	5.6
2	4	6	5	5.3	5.3	5.6
2	4	7	7	6.2	6.2	6.6
2	4	7	7	6.2	6.2	6.6
2	4	7	7	6.2	6.2	6.6
2	4	7	7	6.2	6.2	6.6
2	4	8	8	7.0	7.1	7.5
2	4	8	8	7.0	7.1	7.5
2	4	8	8	7.0	7.1	7.5
2	4	8	8	7.0	7.1	7.5
2	4	8	7	7.0	7.1	7.5
2	4	8	7	7.0	7.1	7.5
2	4	12	12	10.6	10.7	11.3
2	3	1	...	7	7	6.2	5.8	6.4
2	3	1	...	10	4	8.8	8.3	9.1
2	3	1	...	12	12	10.6	10.0	10.9
2	3	...	1	7	6	6.2	5.1	6.0
2	3	...	1	8	8	7.0	5.8	6.9
2	3	...	1	9	9	7.9	6.6	7.7
2	3	...	1	9	9	7.9	6.6	7.7
2	3	...	1	9	7	7.9	6.6	7.7
2	3	...	1	10	10	8.8	7.3	8.6
2	2	2	...	7	7	6.2	5.4	6.2
2	2	2	...	10	9	8.8	7.7	8.8
2	2	1	1	6	6	5.3	4.0	5.0
2	2	1	1	10	10	8.8	6.7	8.3
...	2	...	2	1	1	7	4	6.2	4.7	4.6
...	...	2	2	...	2	8	5	5.4	4.6	4.8
...	...	2	3	...	1	6	2	1.7	4.4	2.2
...	...	2	2	...	2	9	1	2.5	5.1	2.5
...	...	2	1	...	3	6	1	2.7	2.5	1.2
...	...	2	1	...	3	11	3	3.1	4.5	2.2
...	...	2	1	1	2	6	...	1.7	3.0	1.5
...	...	2	1	1	2	7	4	2.0	3.6	1.8
1	1	...	3	1	...	6	6	4.7	5.0	4.9
1	1	...	3	1	...	7	6	5.5	5.7	5.7
1	1	...	3	1	...	8	6	6.2	6.6	6.6
1	1	...	3	1	...	9	7	7.0	7.5	7.4
1	1	...	3	1	...	11	10	8.6	9.1	9.2

Appendix K (continued)

Page 217

TABLE 20—*continued.*

Eye-colours of the						Total child- ren.	Number of the light eye- coloured children.			
Parents.			Grandparents.				Ob- served.	Children.		
Light.	Hazel.	Dark.	Light.	Hazel.	Dark.			I.	II.	III.
1	1	...	3	...	1	9	6	7.0	6.6	6.9
1	1	...	3	...	1	11	7	8.6	8.0	8.5
1	1	...	2	2	...	7	6	5.5	5.4	4.4
1	1	...	2	2	...	9	9	7.0	6.9	5.7
1	1	...	2	2	...	11	1	8.6	8.5	6.9
1	1	...	2	...	2	3	6	4.7	3.4	4.1
1	1	...	2	...	2	6	4	4.7	3.4	4.1
1	1	...	2	...	2	8	5	6.2	4.6	5.5
1	1	...	2	...	2	9	7	7.0	5.1	6.2
1	1	...	2	1	1	6	6	4.7	4.0	4.4
1	1	...	2	1	1	10	9	7.8	6.7	7.4
1	1	...	1	3	...	9	4	7.0	5.5	6.8
1	1	...	1	1	2	8	5	6.2	4.1	5.3
1	...	1	4	7	3	4.1	6.2	4.8
1	...	1	3	...	1	6	4	3.5	4.4	3.7
1	...	1	3	...	1	7	3	4.1	5.1	4.3
1	...	1	3	...	1	8	6	4.6	5.8	4.9
1	...	1	3	...	1	8	5	4.6	5.8	4.9
1	...	1	3	...	1	8	4	4.6	5.8	4.9
1	...	1	3	...	1	9	6	5.2	6.6	5.5
1	...	1	3	...	1	9	5	5.2	6.6	5.5
1	...	1	2	...	2	6	5	3.5	3.4	3.2
1	...	1	2	...	2	6	3	3.5	3.4	3.2
1	...	1	2	...	2	8	4	4.6	4.6	4.2
1	...	1	2	...	2	10	2	5.8	5.7	5.3
1	...	1	2	...	2	14	9	8.1	8.0	7.4
1	...	1	2	1	1	7	5	4.1	4.7	4.1
1	...	1	1	2	1	7	3	4.1	4.3	3.9
1	...	1	1	1	2	7	4	4.1	3.6	3.5
1	...	1	1	...	3	8	4	4.6	3.3	3.6
1	...	1	1	...	3	8	3	4.6	3.3	3.6
1	...	1	...	1	3	6	3	3.5	2.1	2.6
...	1	1	2	...	2	6	3	4.8	3.4	2.6
...	1	1	2	1	1	9	4	7.0	6.0	4.4
...	1	1	1	...	3	13	8	10.1	5.3	4.7
...	1	1	...	4	...	7	2	5.5	4.6	3.4

Appendix L

Calculated Contributions of Eye Color from Different Generations

Francis Galton, *Natural Inheritance* (London: Macmillan and Company, 1889) 213-14.

Page 213

TABLE 17.
CALCULATED CONTRIBUTIONS OF EYE-COLOUR.

Contribution to the heritage from each.	Data limited to the eye-colours of the					
	2 parents.		4 grandparents.		2 parents and 4 grandparents.	
	I.		II.		III.	
	Light.	Dark.	Light.	Dark.	Light.	Dark.
Light-eyed parent.....	0·30	0·25	...
Hazel-eyed parent.....	0·20	0·10	0·16	0·09
Dark-eyed parent	0·30	0·25
Light-eyed grandparent..	0·16	...	0·08	...
Hazel-eyed grandparent.	0·10	0·06	0·05	0·03
Dark-eyed grandparent...	0·16	...	0·08
Residue, rateably assigned	0·28	0·12	0·25	0·11	0·12	0·06

Appendix L (continued)

TABLE 18.
EXAMPLE OF ONE CALCULATION IN EACH OF THE THREE CASES.

Ancestry and their eye-colours.	I.			II.			III.		
	No. about whom data exist.	Contribute to		No. about whom data exist.	Contribute to		No. about whom data exist.	Contribute to	
		Light.	Dark.		Light.	Dark.		Light.	Dark.
Light-eyed parents.	2	0.60
Hazel-eyed parents.	1	0.16	0.09
Dark-eyed parents.	1	...	0.25
Light-eyed grand- parents.....	1	0.16	...	1	0.08	...
Hazel-eyed grand- parents.....	2	0.20	0.12	2	0.10	0.06
Dark-eyed grand- parents..	1	...	0.16	1	...	0.08
Residue, rateably as- signed.....	...	0.28	0.12	...	0.25	0.11	...	0.12	0.06
Total contributions	...	0.88	0.12	...	0.61	0.39	...	0.46	0.54
		1.00			1.00			1.00	

Appendix M

Excerpts of Pearson's Mathematical Arguments in the Homotyposis Paper

Karl Pearson, "Mathematical Contributions to the Theory of Evolution. IX. On the Principle of Homotyposis and Its Relation to Heredity, to the Variability of the Individual, and to that of the Race. Part I. Homotypos in the Vegetable Kingdom," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 197 (1901): 288-91.

Page 288

II. THEORETICAL DISCUSSION OF THE RELATION OF FRATERNAL CORRELATION TO THE CORRELATION OF UNDIFFERENTIATED LIKE ORGANS.

(2.) Let z be the deviation from the mean of the general population of any individual with regard to any character. Let us suppose z to depend upon certain

Page 289

characters in the spermatozoon and certain characters in the ovum from which the individual has developed. These characters cannot of course be determined, still less measured, but we have no reason to doubt their existence. In the particular spermatozoon from which the individual has developed, let them have deviations $x_1, x_2, x_3 \dots$ from their mean values for all the spermatozoa of the race, and let $y_1, y_2, y_3 \dots$ be the corresponding deviations for the ovum characters. Then

$$z = f(x_1, x_2, x_3 \dots y_1, y_2, y_3 \dots)$$

where f is a quite unknown function.

The mean of the z -character will, however, correspond to the mean values of the spermatozoon and ovum characters, and if we suppose the variation of these characters small as compared with their mean value, we assume as usual for such deviations :

$$z = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \dots + \beta_1 y_1 + \beta_2 y_2 + \beta_3 y_3 + \dots \quad (i.),$$

where the α 's and β 's are independent of the x 's and y 's, and define the male and female inheritance.

Now let σ be the standard deviation of the character z in the population; σ_q the standard deviation of x_q , σ'_q of y_q . Let r_{pq} be the correlation of x_p and x_q , r'_{pq} of y_p and y_q . Then we will suppose that there is no selection of particular ova by particular spermatozoa, or that x_p and y_q are not correlated. Then if $n =$ number of individuals in the population :

$$\sigma^2 = \frac{S(z^2)}{n} = \Sigma \alpha_p^2 \frac{S(x_p^2)}{n} + 2\Sigma(\alpha_p \alpha_q) \frac{S(x_p x_q)}{n} + \Sigma \beta_p^2 \frac{S(y_p^2)}{n} + 2\Sigma(\beta_p \beta_q) \frac{S(y_p y_q)}{n},$$

where S is the sum for all individuals of any x or y for constant subscript, and Σ is the sum of α and β for every possible subscript. This follows by simple squaring and remembering that $S(x_p y_q) = 0$. We thus reach :

$$\sigma^2 = \Sigma(\alpha_p^2 \sigma_p^2) + \Sigma(\beta_p^2 \sigma'_p{}^2) + 2\Sigma(\alpha_p \alpha_q \sigma_p \sigma_q) + 2\Sigma(\beta_p \beta_q \sigma'_p \sigma'_q). \quad \text{(ii.)}$$

Now let us consider the correlation of two individuals due to the spermatozoa and ova put forth by the same two individuals. Let z_1 and z_2 be the values of their characters, and x', x'', y', y'' represent the fundamental characters in the two spermatozoa, and two ova on which they depend.

Then we have

$$z_1 = \Sigma(\alpha_p x'_p) + \Sigma(\beta_p y'_p) \\ z_2 = \Sigma(\alpha_p x''_p) + \Sigma(\beta_p y''_p).$$

Now let us multiply z_1 by z_2 and sum for every fraternal pair; then if R be the

Page 290

fraternal correlation we have, since one set of brethren will have the same standard deviation as the other, if they be taken of the same sex,

$$S(z_1 \times z_2) = n \times \sigma \times \sigma \times R,$$

where n is the total number of pairs of brethren.

Hence

$$n \times \sigma^2 \times R = \Sigma(\alpha_p^2 (x'_p x''_p)) + \Sigma(\alpha_p \alpha_q S(x'_p x''_q + x''_p x'_q)) \\ + \Sigma(\beta_p^2 S(y'_p y''_p)) + \Sigma(\beta_p \beta_q S(y'_p y''_q + y''_p y'_q))$$

with the same notation as before and the same assumption as to no selective process between the female and male reproductive cells.

Now x'_p, x''_p are undifferentiated like organs put forth by an individual. If ρ_{pp} be the correlation of such :

$$S(x'_p x''_p) = n \sigma_p^2 \rho_{pp},$$

Similarly

$$S(y'_p y''_p) = n \sigma'_p{}^2 \rho'_{pp}.$$

In the next place we should have *all* pairs of brothers :

$$S(x'_p x''_q) = S(x''_p x'_q) = \sigma_p \sigma_q \rho_{pq},$$

$$S(y'_p y''_q) = S(y''_p y'_q) = \sigma'_p \sigma'_q \rho'_{pq}.$$

We thus deduce

$$\sigma^2 \times R = \Sigma(\alpha_p^2 \sigma_p^2 \rho_{pp}) + \Sigma(\beta_p^2 \sigma'_p{}^2 \rho'_{pp}) \\ + 2\Sigma(\alpha_p \alpha_q \sigma_p \sigma_q \rho_{pq}) + 2\Sigma(\beta_p \beta_q \sigma'_p \sigma'_q \rho'_{pq}). \quad \text{(iii.)}$$

Now if the degree of resemblance between undifferentiated like organs vary largely from organ to organ, we could proceed no further with this investigation. We should have shown that there was a relation, namely (iii.) between the R and the ρ 's, but as the characters in the spermatozoa and ova are unknown and unmeasurable, we could not proceed further. On the other hand, if the correlation between undifferentiated like organs is approximately constant, then we may replace correlation coefficients like ρ_{pp} and ρ'_{pq} by their *mean* value ρ , say.

We have next to consider such correlations as ρ_{pq} or ρ'_{pq} . These are the two correlations between two different characters in two spermatozoa or ova from the same individual. Such correlations must vanish (i.) if there be no organic correlation between characters of the same spermatozoon or ovum, (ii.) if there be no correlation between the reproductive cells put forth by the same individual. In other words, ρ_{pq} must contain ρ and r_{pq} as factors, and since if these two correlations are perfect ρ_{pq} must be perfect, it is a reasonable hypothesis to assume

$$\rho_{pq} = \rho \times r_{pq} \dots \dots \dots (iv.).$$

This, it will be seen, is identical in form with the result I have supposed to hold

Page 291

for cross-heredity,* and which appears to be approximately true for that case. We shall consider later statistics bearing on this result.

Assuming accordingly that (iv.) holds, we find from (iii.)

$$\begin{aligned} \sigma^2 \times R &= \rho \{ \Sigma (\alpha_p^2 \sigma_p^2) + \Sigma (\beta_p^2 \sigma'_p{}^2) + 2\Sigma (\alpha_p \alpha_q \sigma_p \sigma_q r_{pq}) + 2\Sigma (\beta_p \beta_q \sigma'_p \sigma'_q r'_{pq}) \} \\ &= \rho \times \sigma^2 \text{ by (ii.).} \end{aligned}$$

Hence we conclude that

$$R = \rho \dots \dots \dots (v.).$$

Or: *The correlation between brothers will be equal to the mean correlation between the undifferentiated like organs put forth by an individual.*

Now, if this result be true, it is very remarkable and very fundamental. We should hardly expect it to be absolutely true, for it is very unlikely that the coefficient of correlation between undifferentiated like organs is the same whatever the organs may be. This equality may rather be spoken of as belonging to an ideal theoretical vital state approaching the actual state, perhaps, as BOYLE'S Law or the perfect fluid approach phenomena observed in physical nature. What we should expect would be a general approximation between the values of R and ρ , and a tendency to equality when large series are averaged. This is the point which we shall investigate in the sixth section of this paper, after placing before the reader in the fourth section a fairly wide† range of actual statistics.

Appendix N

Pearson's Table of Average Homotypic Correlations

Karl Pearson, "Mathematical Contributions to the Theory of Evolution. IX. On the Principle of Homotyposis and Its Relation to Heredity, to the Variability of the Individual, and to that of the Race. Part I. Homotypos in the Vegetable Kingdom," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 197 (1901): 356-57.

Page 356

TABLE XXXII.—General Results for Homotypic Correlation.

Race.	Character.	Per-centage variation.	Corre-lation.	Remarks.
Mushroom, Hampden . . .	Lengths of gills .	50·92	·8607	} All these results introduce a correlation due to stages of growth and accordingly are not included in the determination of means.
" " . . .	Breadths of gills .	67·67	·7363	
" " . . .	Lengths and breadths of gills	—	·6275	
Wild Ivy, mixed localities .	Lengths of leaves .	82·73	·5618	
" " . . .	Breadths of leaves	84·56	·5332	
" " . . .	Lengths and breadths of leaves	—	·5157	
(i.) Ceterach, Somersetshire	Lobes on fronds .	77·57	·6311	} Said to be largely affected by growth and environment.
(ii.) Hartstongue, Somersetshire	Sori on fronds .	77·64	·6303	
(iii.) Shirley Poppy, Chelsea .	Stigmatic bands .	78·86	·6149	} Much selected in transit. Possibly slightly heterogeneous.
(iv.) English Onion, Hampden	Veins in tunics .	79·18	·6108	
(v.) Holly, Dorsetshire . . .	Prickles on leaves	80·11	·5985	} Heterogeneous.
(vi.) Spanish Chestnut, mixed	Veins in leaves .	80·65	·5913	
(vii.) Beech, Buckinghamshire	Veins in leaves .	82·17	·5699	} Possibly influenced by individual growth.
(viii.) <i>Papaver Rhoeas</i> , Hampden	Stigmatic bands .	82·71	·5620	
(ix.) Mushroom, Hampden .	Gill indices . . .	83·58	·5490	
(x.) <i>Papaver Rhoeas</i> , Quantocks	Stigmatic bands .	84·59	·5333	} All from one field.
(xi.) Shirley Poppy, Hampden	Stigmatic bands .	85·18	·5238	
(xii.) Spanish Chestnut, Buckinghamshire	Veins in leaves .	88·51	·4655	
(xiii.) Broom, Yorkshire . . .	Seeds in pods . .	90·96	·4155	} From two localities and possibly slightly influenced by differentiation.
(xiv.) Ash, Monmouthshire . .	Leaflets on leaves .	91·44	·4047	
(xv.) <i>Papaver Rhoeas</i> , Lower Chilterns	Stigmatic bands .	91·66	·3997	
(xvi.) Ash, Dorsetshire . . .	Leaflets on leaves .	91·81	·3964	
(xvii.) Ash, Buckinghamshire .	Leaflets on leaves .	92·73	·3743	
(xviii.) Holly, Somersetshire .	Prickles on leaves	93·12	·3548	
(xix.) Wild Ivy, mixed localities	Leaf indices . . .	96·21	·2726	} Differentiation of organs due to position on stem.
(xx.) <i>Nigella Hispanica</i> , Slough	Segments of seed-capsules	98·18	·1899	
(xxi.) <i>Malva Rotundifolia</i> , Hampden	Segments of seed-vessels	98·32	·1827	} Principally spread from one clump by stolons.
(xxii.) Woodruff, Buckinghamshire	Members of whorls	98·49	·1733	
Mean of 22 cases . . .	—	87·44	·4570	—

Appendix N (continued)

Pearson's Table of Average Fraternal Correlations

Page 357

TABLE XXXIII.—General Results for Fraternal Correlation.*

Race.	Sex.	Character.	Source of material.	No. of cases.	Reduced by	Correlation.	Remarks.
(i.) Daphnia	♀ & ♀	Length of spine	ERNEST WARREN	330	K. PEARSON6934	Probably much too high, owing to heterogeneity introduced by the selection of a few mothers only.
(ii.) Horse	♀ & ♀	Coat-colour.	WEATHERBY'S Studbooks .	1000	K. PEARSON, L. BRAMLEY-MOORE, and A. LEE	.6928	Probably much too high, owing to heterogeneity introduced by the use of comparatively few sires.
(iii.) "	♂ & ♂	"	"	1000		.6232	
(iv.) "	♂ & ♀	"	"	1000		.5827	
(v.) Man.	♀ & ♀	Forearm.	PEARSON, family data . . .	441	A. LEE5424	One pair only from each family.
(vi.) Hound (Basset)	mixed	Coat-colour.	GALTON, from studbook . .	—	K. PEARSON and A. LEE5257	All members of litter without regard to sex.
(vii.) Man.	♂ & ♂	Eye-colour.	GALTON, family data	1500	K. PEARSON5169	All possibly pairs in family taken.
(viii.) "	♀ & ♀	Cephalic index	FRANZ BOAS, N. A. Indians	—	C. FAWCETT4890	Paternity doubtful.
(ix.) "	♂ & ♀	Eye-colour.	GALTON, family data	1500	K. PEARSON4615	See remark to (vii.).
(x.) "	♀ & ♀	"	"	1500	"	.4463	"
(xi.) "	♀ & ♀	Stature	"	595	"	.4436	"
(xii.) "	♂ & ♂	"	"	605	"	.3913	"
(xiii.) "	♂ & ♂	Cephalic index	FRANZ BOAS, N. A. Indians	—	C. FAWCETT3790	See remark to (viii.).
(xiv.) "	♂ & ♀	Stature	GALTON, family data	1181	K. PEARSON3754	See remark to (vii.).
(xv.) "	♂ & ♀	Cephalic index	FRANZ BOAS, N. A. Indians	—	C. FAWCETT3400	See remark to (viii.).
(xvi.) "	♀ & ♀	Longevity	Quaker records	1050	M. BEETON3323	Reduced below true value by non-selective deaths.
(xvii.) "	mixed	Temper	GALTON, family data	1294	K. PEARSON3167	Character very indefinite, and difficult to estimate.
(xviii.) "	♂ & ♂	Longevity	Peerage records	1000	M. BEETON2602	See remark to (xvi.).
(xix.) "	♂ & ♀	"	Quaker records	1947	"	.1973	"
—	—	—	—	—	Mean of 19 series4479	—

* [Since the above memoir was written I have deigned another exceedingly interesting value for fraternal correlation from the measurements of Prof. C. B. DAVENPORT on the statoblasts of the Bryozoa (*Pectinella Magnifica*, LEBRY). See 'The American Naturalist', vol. 34, p. 964, 1900. DAVENPORT gives for the standard-deviation of the number of hooks in all statoblasts the value 1.336, and for the average standard deviation of colonies of statoblasts 1.197. If we represent the former by σ , the latter will be $\sigma\sqrt{1-p^2}$, whence I find for the fraternal correlation $r = .4302$, a result in excellent agreement with the mean values we have just found.—July, 1901.]

References

- Aristotle. *The Rhetoric and the Poetics of Aristotle*. Trans. W. Rhys Roberts and Ingram Bywater. New York: Modern Library College Editions, 1984.
- Bateson, William. "Heredity, Differentiation, and other Conceptions of Biology: A Consideration of Professor Karl Pearson's Paper 'On the Principle of Homotyposis.'" *Proceedings of the Royal Society of London* 69 (1901-02): 193-205.
- Bateson, William. *Materials for the Study of Variation*. London: Macmillan and Co., 1894.
- Bateson, William. *Mendel's Principles of Heredity: A Defense*. Cambridge: Cambridge UP, 1902.
- Bateson, William. "Some Cases of Variation in Secondary Sexual Characters Statistically Examined." *Proceedings of the Royal Society* 54 (1893): 585-94.
- Bowlby, John. *Charles Darwin a New Life*. New York: Norton, 1991.
- Bowler, Peter. *Charles Darwin the Man and his Influence*. Cambridge: Cambridge UP, 1990.
- Brent, Peter. *Charles Darwin a Man of Enlarged Curiosity*. New York: Harper and Row, 1981.
- Browne, Janet. *Charles Darwin the Power of Place*. Vol. 2. Princeton: Princeton UP, 1996.
- Browne, Janet. "Darwin's Botanical Arithmetic and the 'Principle of Divergence,' 1854-1858." *Journal of the History of Biology* 13 (1980): 53-89.
- Brookes, Martin. *Extreme Measures*. New York: Bloomsbury, 2004.
- Bulmer, Michael. *Francis Galton Pioneer of Heredity and Biometry*. Baltimore: Johns Hopkins UP, 2003.
- Campbell, John Angus. "Charles Darwin Rhetorician of Science." *The Rhetoric of the Human Sciences Language and Argument in Scholarship and Public Affairs*. Ed. John S. Nelson. Madison: The U of Wisconsin P, 1987.
- Campbell, John Angus. "The Invisible Rhetorician: Charles Darwin's 'Third Party' Strategy." *Rhetorica* 7.1 (1989): 55-85.

- Campbell, John Angus. "Why was Darwin Believed?: Darwin's *Origin* and the problem of Intellectual Revolution." *Configurations* 11.2 (2003): 203-37.
- Cannon, Susan. *Science in Culture: The Early Victorian Period*. New York: Dawson, 1978.
- Chatterjee, Shoutir Kishore. *Statistical Thought: A Perspective and History*. Oxford: Oxford UP, 2003.
- "Combinatorics" and "combinatorial." *Merriam-Webster's Collegiate Dictionary*. 10th ed. 1998.
- Corcos, Alain; and Floyd Monaghan. *Gregor Mendel's Experiments on Plant Hybrids*. New Brunswick, New Jersey: Rutgers UP, 1966.
- Cullen, Michael. *The Statistical Movement in Early Victorian Britain*. New York: The Harvester Press, 1975.
- Darwin, Charles. *Autobiographies*. Michael Neve and Sharon Messenger, eds. London: Penguin Books. 1986.
- Darwin, Charles. *Charles Darwin's Natural Selection*. Ed. R.C. Stauffer. Cambridge: Cambridge UP, 1975.
- Darwin, Charles. *The Origin of Species*. Oxford: Oxford University Press, 1996.
- Darwin, Charles. *The Variation of Animals and Plants under Domestication*. 2nd ed. 2 vols. New York: D. Appleton & Co., 1883.
- Darwin, F. and Seward, A.C. eds. *More Letters of Charles Darwin*. 2 vols. London: John Murray, 1903.
- Daston, Lorraine. *Classical Probability in the Enlightenment*. Princeton: Princeton UP, 1988.
- Dewey, John. "Galton's Statistical Methods." *Publications of the American Statistical Association* 1 (1889): 331-34.
- Fahnestock, Jeanne. *Rhetorical Figures in Science*. New York: Oxford UP, 1999.
- Gale, Berry G. *Evolution without Evidence Charles Darwin and the Origin of Species*. Albuquerque: U of New Mexico P, 1982.
- Galton, Francis. *Hereditary Genius*. New York: St. Martin's Press, 1978.

- Galton, Francis. *Memories of My Life*. New York: E.P. Dutton and Company, 1909.
- Galton, Francis. *Natural Inheritance*. London: Macmillan and Company, 1889.
- Galton, Francis. "Regression Towards Mediocrity in Hereditary Stature." *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 15 (1886): 246-63.
- "Gene." *Oxford English Dictionary*. Oxford: Oxford UP, 2006. 27 Mar. 2006 <<http://dictionary.oed.com/>>.
- Ghislen, Michael T. *The Triumph of the Darwinian Method*. Berkeley: U of California P, 1969.
- Gillham, Nicholas Wright. *A Life of Sir Francis Galton*. Oxford: Oxford UP, 2001.
- Gillespie, Neal C. *Charles Darwin and the Problem of Creation*. Chicago: U of Chicago P, 1979.
- Gillispie, Charles. *Pierre Simon Laplace 1749 – 1827: A Life in Exact Science*. Princeton: Princeton UP, 1997.
- Gonick, Larry and Woolcott Smith. *The Cartoon Guide to Statistics*. New York: Harper Collins, 1993.
- Gould, Stephen J. "False Premise Good Science." *Natural History* 92.10 (1983): 20-26.
- Hacking, Ian. *The Taming of Chance*. Cambridge: Cambridge UP, 1990.
- Hahn, Roger. *Pierre Simon Laplace 1749-1827: A Determined Scientist*. Cambridge: Harvard UP, 2005.
- Hald, Andreas. *A History of Probability and Statistics and Their Applications Before 1750*. Hoboken, NJ: John Wiley and Sons, 2003.
- Hall, Tord. *Carl Friedrich Gauss: A Biography*. Trans. Albert Froderberg. Cambridge, MA: The MIT Press, 1970.
- Henig, Robin Marantz. *The Monk in the Garden*. Boston: Houghton Mifflin, 2000.
- "Heterozygote." *Oxford English Dictionary*. Oxford: Oxford UP, 2006. 27 Mar. 2006 <<http://dictionary.oed.com/>>.
- Hull, David. *Darwin and His Critics: The reception of Darwin's Theory of Evolution by the Scientific Community*. Cambridge: Harvard UP, 1973.

- Iltis, Hugo. *Life of Mendel*. Trans. Eden and Cedar Paul. New York: W.W. Norton & Company, 1932.
- Kant, Immanuel. *Metaphysical Foundations of Natural Science*. Ed. Michael Friedman. Cambridge: Cambridge UP, 2004.
- Kölreuter, Joseph Gottlieb. *Vorläufige Nachricht von einigen das Geschlecht der Pflanzen betreffenden Versuchen und Beobachtungen*. Microform. Leipzig: Gleditschischen Handlung, 1761-1766.
- Laplace, Pierre-Simon. *Philosophical Essays on Probabilities*. Trans. Andrew Dale. New York: Springer-Verlag, 1995.
- “Mathematics.” *Merriam Webster’s Collegiate Dictionary*. 10th ed. 1998.
- Magnello, Eileen M. “Karl Pearson’s Gresham Lectures: W.F.R. Weldon, Speciation and the Origins of Pearsonian Statistics.” *British Journal for the History of Science* 29 (1996): 43-63.
- Malthus, T.R. *An Essay on the Principle of Population*. 1826. London: Ward, Lock, and Co., 1890.
- Malthus, T.R. *An Essay on the Principle of Population*. Cambridge: Cambridge UP, 1992.
- Manier, Edward. “‘External Factors’ and ‘Ideology’ in the Earliest Drafts of Darwin’s Theory.” *Social Studies of Science* 17.4 (1987): 581-609.
- Mayr, Ernst. *Charles Darwin and the Genesis of Modern Evolutionary Thought*. Cambridge: Harvard UP, 1991.
- Mayr, Ernst. *The Growth of Biological Thought*. Cambridge: Harvard UP, 1982.
- Mayr, Ernst. *One Long Argument*. Harvard: Harvard UP, 1991.
- Mendel, Gregor. *Experiments in Plant Hybridization*. Cambridge: Harvard UP, 1965.
- Mendel, Gregor. *Versuche über Pflanzen-Hybriden*. J. Cramer and H.K. Swann eds. New York: Hafner Publishing, 1960.
- Moore, Randy. “The Persuasive Mr. Darwin.” *Bioscience* 47.2 (1997): 107-14.
- Nägeli, Karl. “Die Bastardbildung im Pflanzenreich.” *Botanische Mittheilungen* 2 (1865): 187-235.
- “Natural Inheritance.” *Science* 13 (1889): 266-67.

- Olby, Robert. "The Dimensions of Scientific Controversy: The Biometric-Mendelian Debate." *British Journal for the History of Science* 22 (1988): 299-320.
- Olby, Robert. *Origins of Mendelism*. New York: Schocken Books, 1966.
- Orel, Vitezslav. *Mendel*. Oxford: Oxford UP, 1984.
- Pearson, Karl. *The Grammar of Science*. London: Charles Scribner's Sons, 1892.
- Pearson, Karl. "Mathematical Contributions to the Theory of Evolution." *Proceedings of the Royal Society London* 45 (1893): 329-33.
- Pearson, Karl. "Mathematical Contributions to the Theory of Evolution. IX. On the Principle of Homotyposis and Its Relation to Heredity, to the Variability of the Individual, and to that of the Race. Part I. Homotypos in the Vegetable Kingdom." *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 197 (1901): 285-379.
- Pearson, Karl. "The Scope of Biometrika." *Biometrika* 1.1 (1901): 1-2.
- Pearson, Karl. "The Spirit of Biometrika." *Biometrika* 1.1 (1901): 3-6.
- Porter, Theodore M. *Karl Pearson*. Princeton: Princeton University Press, 2004.
- Porter, Theodore M. *The Rise of Statistical Thinking 1820-1900*. Princeton: Princeton UP, 1986.
- Provine, William. *The Origins of Theoretical Population Genetics*. Chicago: U of Chicago P, 1971.
- Punnett, R.C., ed. *Scientific Papers of William Bateson*. 2 vols. Cambridge: Cambridge UP, 1928.
- Quetelet, Adolphe. *A Treatise on Man and the Development of his Faculties*. Gainesville, Florida: Scholar's Facsimiles & Reprints, 1969.
- Roberts, H.F. *Hybridization Before Mendel*. Princeton: Princeton UP, 1929.
- Shull, Geo H. "Heredity as an Exact Science." *Botanical Gazette* 50.3 (1910): 226-29.
- Stern, Curt and Eva Sherwood eds. *The Origin of Genetics a Mendel Source Book*. San Francisco: W.H. Freeman and Company, 1966.
- Venn, John. "Natural Inheritance." *Mind* 14.55 (1889): 414-20.

Weldon, Frank. "On Certain Correlated Variations in *Carcinus moenas*."
Proceedings of the Royal Society 54 (1893): 585-94.

Weldon, W.F.R. "Remarks on Variations in Plants and Animals. To Accompany the First Report of the Committee for Conducting Statistical Inquiries into the Measurable Characteristics of Plants and Animals." *Proceedings of the Royal Society of London* 57 (1894-5): 379-82.

Works Consulted

- Aydon Cyril. *Charles Darwin: The Naturalist Who Started a Scientific Revolution*. New York: Carroll and Graf Publishers, 2002.
- Bowler, Peter. *The Mendelian Revolution*. Baltimore: The Johns Hopkins UP, 1989.
- Cerrai, Paola; Paola Freguglia; and Claudio Pellegrini. *The Application of Mathematics to the Sciences of Nature*. New York: Plenum Publishers, 2002.
- Cicero, Marcus Tullius. *Topica*. Ed. Tobias Reinhardt. Oxford: Oxford UP, 2003.
- Cock, A.G. "William Bateson, Mendelism and Biometry." *Journal of the History of Biology* 6.1 (1973): 1-36.
- Copi, Irving M. and Carl Cohen. *Essentials of Logic*. Upper Saddle River, NJ: Pearson, 2004.
- Darden, Lindley. *Theory Change in Science*. New York: Oxford UP, 1991.
- Darden, Lindley. "William Bateson and the Promise of Mendelism." *Journal of the History of Biology* 10.1 (1977): 87-106.
- Darwin, F. ed. *The Life and Letters of Charles Darwin*. New York: D. Appleton and Company, 1905.
- David, H.A. and A.W.F. Edwards. *Annotated Readings in the History of Statistics*. New York: Springer-Verlag, 2001.
- Davis, Philip and Reuben Hersh. *The Mathematical Experience*. Boston: Houghton Mifflin, 1981.
- Davis, Philip and Reuben Hersh. "Rhetoric and Mathematics." *The Rhetoric of the Human Sciences Language and Argument in Scholarship and Public Affairs*. Ed. John S. Nelson. Madison: The U of Wisconsin P, 1987.
- Falk, Raphael. "The Struggle of Genetics for Independence." *Journal of the History of Biology* 28.2 (1995): 219-46.
- Farral, Lyndsay. "Controversy and Conflict in Science: A Case Study—the English Biometric School and Mendel's Laws." *Social Studies of Science* 5.3 (1975): 269-301.
- Fisher, R.A. "Has Mendel's Work been Rediscovered?" *Annals of Science* 1 (1936): 115-37.

- Forsdyke, Donald R. "William Bateson." *Encyclopedia of Life Sciences*. Vol. 3. New York: Macmillan Publishers, 2002.
- Galton, Francis. "Co-relations and their Measurement, Chiefly from Anthropometric Data." *Proceedings of the Royal Society London* 45 (1888): 135-45.
- Galton, Francis and J.D. Hamilton Dickenson. "Family Likeness in Stature." *Proceedings of the Royal Society London* 40 (1886): 42-73.
- Heidelberger, Michael, Lorenz Kruger, and Rosemarie Rheinwald, eds. *Probability Since 1800*. Bielefeld, Germany: U of Bielefeld, 1983.
- Kline, Morris. *Mathematical Thought From Ancient to Modern*. New York: Oxford UP, 1972.
- Kolmogorov, A.N., and A.P. Yushkevich. *Mathematics of the 19th Century*. Basel, Switzerland: Birkhauser Verlag, 1992.
- Koosis, Donald K. *Statistics a Self Teaching Guide*. New York: John Wiley & Sons, 1997.
- Laplace, Pierre-Simon. *Celestial Mechanics*. Vol. 3 Bronx, NY: Chelsea Publishing Company, 1966-69.
- Lindberg, David C. *The Beginnings of Western Science*. Chicago: U of Chicago P, 1992.
- Magnello, Eileen M. "Karl Pearson's Mathematization of Inheritance: From Ancestral Heredity to Mendelian Genetics (1895-1909)." *Annals of Science* 55 (1998): 35-94.
- Magner, Lois N. *A History of the Life Sciences*. New York: Marcel Dekker, 1994.
- Maistrov, L.E. *Probability Theory: A Historical Sketch*. New York: Academic Press, 1974.
- McMurray, Emily J. ed. "William Bateson." *Notable Twentieth-Century Scientists*. Vol.1. New York: International Thomson Publishing, 1995.
- Norton, B.J. "The Biometric Defense of Darwinism." *Journal of the History of Biology* 6.2 (1973): 283-316.
- Olbrechts-Tyteca, L. Chiam Perelman. *The New Rhetoric*. Notre Dame: U of Notre Dame P, 1971.
- Olby, Robert. "Mendelism: From Hybrids and Trade to Science." *Comptes rendus de l'Académie des sciences. Série III, Sciences de la vie* 323.12 (2000): 1043-51.

- Olby, Robert. "William Bateson's Introduction of Mendelism to England: A Reassessment." *British Journal for the History of Science* 20 (1987): 399-420.
- Olby, Robert. "William Bateson." *Oxford Dictionary of National Biography*. Vol.4 Eds. H.C.G. Matthew and Brian Harrison. Oxford: Oxford UP, 2004.
- Owen, D.B. ed. *On the History of Statistics and Probability*. New York: Marcel Dekker, 1976.
- Pearson, Karl. "On the Fundamental Concepts of Biology." *Biometrika* 1.3 (1902): 320-44.
- Pearson, Karl. *The Grammar of Science*. London: Adam and Charles Black, 1900.
- Pearson, Karl. *The Grammar of Science*. London: J.M. Dent & Sons, 1937.
- Pearson, Karl. *The Grammar of Science*. Gloucester, Mass.: Peter Smith, 1969.
- Piegorsch, Walter W. "The Gregor Mendel Controversy." *History of Science* 24 (1986): 173-82.
- Porter, Theodor M. "Statistics and Physical Theories." *The Cambridge History of Science*. Vol. 5 Ed. Mary Jo Nye. Cambridge: Cambridge UP, 2003.
- Root-Bernstein, Robert S. "Mendel and Methodology." *History of Science* 21 (1983): 275-95.
- Rushton, Alan R. "Nettleship, Pearson and Bateson: The Biometric-Mendelian Debate in a Medical Context." *Journal of the History of Medicine* 55.2 (2000): 134-57.
- Schlagel, Richard. *From Myth to Modern Mind*. 2 vols. New York: Peter Lang, 1995.
- Sloan, Phillip R. "Mach's Phenomenalism and the British reception of Mendelism." *Comptes rendus de l'Académie des sciences. Série III, Sciences de la vie* 323.12 (2000): 1069-79.
- Svitak, Sylvia. "Statistics and Probability Theory." *History of Modern Science and Mathematics*. Ed. Brian Baigrie. Vol. 4. New York: Charles Scribner's Sons, 2002.
- Waller, John. "Gentlemanly Men of Science: Sir Francis Galton and the Professionalization of the British Life Sciences." *Journal of the History of Biology* 34 (2001): 83-114.

Weldon, W.F.R. "An Attempt to Measure the Death Rate Due to the Selective Destruction of *Carcinus moenas* with Respect to a Particular Dimension." *Proceedings of the Royal Society London* 57 (1894-95): 360-79.

Weldon, W.F.R. "Mendel's Laws of Alternative Inheritance in Peas." *Biometrika* 1.2 (1902): 228-54.

Weldon, W.F.R. "The Variations Occurring in Certain Decapod Crustaceans. I. *Cragon vulgaris*." *Proceedings of the Royal Society* 47 (1889-90): 445-53.