

Research Article

Erich Benjamin Pracht*, Thomas McCauley

Paul's Style and the Problem of the Pastoral Letters: Assessing Statistical Models of Description and Inference

<https://doi.org/10.1515/opth-2024-0034>

received September 26, 2024; accepted November 28, 2024

Abstract: In this study, we combine the core competencies of a New Testament scholar and a mathematician to conduct a stylometric analysis of the Pauline and Pastoral Letters based on 18 epistolary modes of discourse, which we argue are at least as stylistically relevant as traditional features such as term frequencies. By focusing on features of this type, we apply discrete probability models – namely, the binomial, Poisson, and negative binomial distributions – to determine which ones are the best fit for each feature. With our approach, we have found a way to utilize Paul's authentic letters as a standard of measurement by which the style(s) of the Pastorals can be compared, an essential first step to statistical inference that is generally lacking in Pauline stylometry. With few exceptions, we find that the Pastoral Letters do not contain stylistic differences with the authentic Paulines at a statistically significant level, a finding that disputes many previous studies (especially in the qualitative tradition) that assume stylistic difference, but which is in line with an emerging trend in computational analyses of the Pauline and Pastoral Letters.

Keywords: Pastoral Letters, Paul, style, stylometry, authorship

1 Introduction

In New Testament studies, the style(s) of the Pastoral Letters in relation to the other letters in the Pauline corpus has long been a *crux interpretum*. The issue of style¹ is discussed at length in the commentaries and other qualitative studies, which constitutes most scholarly production in biblical studies. Nevertheless, although the digital humanities generally keep a low profile in New Testament studies, there has been a proliferation of statistical and/or computational approaches to the problem (Section 2). Stylometry – that is, the quantitative analysis of style – has uncovered a range of insights into the problem of Pastoral Letters. We contend, however, that two major desiderata remain in stylometric studies of the Pauline and Pastoral Letters:

1. The field lacks clarity on how to use the Pauline letters as a standard of measurement.² Virtually all scholars who work on the problem make judgments about the authenticity of the Pastorals by way of comparison to a group of letters deemed authentically Pauline. Although this premise is theoretically sound, sufficient effort has not been made to assess mathematically the defining features of Paul's style.

¹ For our understanding of style, see Section 3.

² To our knowledge, the two exceptions are Mealand, "The Extent of the Pauline Corpus;" and Robertson, *Digital Approaches to Paul's Letters*.

* **Corresponding author: Erich Benjamin Pracht**, Department of Theology, Aarhus University, Aarhus C, Denmark, e-mail: ebp@cas.au.dk

Thomas McCauley: Department of Mathematics, Arlington High School, Arlington, MA, United States

2. The field over-analyzes the same types of stylistic features. Up until now, the field almost exclusively operates within the framework of traditional stylometry. In other words, units of analysis tend to be limited to lexical and syntactical features.³ By contrast, we define style in terms of modes of epistolary discourse and thereby consider features outside of the bounds of traditional stylometry, which we argue is essential to moving the discussion of the Pauline and Pastoral Letters forward.

This article will unfold as follows: to explain how we arrived at the two desiderata listed above, we provide a literature review of statistical approaches to the Pauline and Pastoral Letters (Section 2); next, we discuss our process of feature selection with an eye toward moving beyond lexical and syntactical elements of style (Section 3); next, we assess how well 18 stylistic features in Paul's authentic letters fit various probability models, the purpose of which is to determine how to use Paul's letters as a standard of measurement (Section 4); finally, we compare Paul's style to the style(s) of the Pastoral Letters using the probability models that represent the best fit (Section 5).

2 Literature Review

2.1 The Debate on Paul and the Pastorals

In this section, we review the major trends in the history of the debate on the authenticity of the Pastoral Letters with a particular focus on the role of style and computational methodologies. Up until the introduction of historical-critical methods in Enlightenment-era biblical scholarship, New Testament interpreters have assumed the Pauline authorship of the letters to Timothy and Titus.⁴ In the nineteenth century, following initial challenges to the Pauline authorship of the Pastorals, some New Testament scholars defended their authenticity.⁵ In the twentieth century, the view that Paul did not write the Pastorals gained traction, but some critical scholars nevertheless produced commentaries on the Pastorals from the perspective of Pauline authorship.⁶ In the early twenty-first century, the majority of scholars think that the Pastoral Letters were not written by Paul, although there is the occasional critical argument to the contrary.⁷

Scholars who adjudicate the debate operate, as far as we can tell, with the unstated yet reasonable premise that Paul's literary output provides a standard of measurement against which the Pastorals can be compared. In other words, there is in the New Testament a group of "authentic" letters that can serve as a touchstone.⁸ In connection with this premise, debates on the Pauline and Pastoral Letters generally revolve around the following issues:

- The Pastorals envision a structured church with well-defined roles for local functionaries beyond anything seen in the authentic Pauline letters. In this connection, the Pastoral Letters lack the degree of eschatological urgency found in the authentic Paulines, which suggests that these letters were written in a context when

³ To our knowledge, the major exception is Robertson, *Paul's Letters and Contemporary Greco-Roman Literature*. For discussion of how Robertson's work influences our approach, refer to Section 3.

⁴ For a more detailed review, refer to van Nes, *Pauline Language and the Pastoral Epistles*, 5–36.

⁵ For example, Koehling, *Der Erste Briefe Pauli an Timeotheus*; Kühl, *Die Gemeindeordnung in den Pastoralbriefen*; Torm, "Über die Sprache in den Pastoralbriefen."

⁶ For example, Spicq, *Les Épîtres Pastorales*; Kelly, *A Commentary on the Pastoral Epistles*; Jeremias, *Die Briefe an Timotheus und Titus*; Holtz, *Die Pastoralbriefe*; Knight, *The Pastoral Letters*; Mounce, *Pastoral Epistles*; Johnson, *The First and Second Letters to Timothy*; see also Prior, *Paul the Letter-Writer and the Second Letter to Timothy*, who makes the case for the Pauline authorship of 2 Timothy.

⁷ Most recently Porter, *The Pastoral Epistles*.

⁸ Throughout the history of criticism, scholars have lodged different views on which letters are "authentic." Today, the majority opinion is that the "authentic" Pauline letters are Romans, 1 and 2 Corinthians, Galatians, 1 Thessalonians, Philippians, and Philemon.

Christians recognized that they must formally organize community life. Also in this connection, some scholars view the Pastorals as responding to an early form of Gnosticism.⁹

- The Pastorals have a different Christology in comparison to the rest of the Pauline corpus. The Pastorals describe Jesus largely in terms of a “savior” and speak of his “appearance,” both of which are uncharacteristic of the authentic Paulines. Other theological differences include the understanding of faith, which in the Pastorals tends to denote a body of religious teachings, whereas in Paul’s authentic letters, faith is most often described in terms of a religious disposition. Furthermore, the emphasis in the Pastorals on the importance of good doctrine and sound teachings also figures heavily in the debate.¹⁰
- The Pastorals use ethical terms and metaphors that differ from what is found in the authentic Paulines. Familiar Pauline concepts related to one’s relationship with God, such as love and one’s existence “in Christ,” take on different meanings in the Pastorals. Furthermore, in the Pastorals, there is an emphasis on “conscience,” which is not heavily thematized in the authentic Paulines.¹¹
- The Pastorals advance a highly restricted view of the role of women in the assembly of believers and in general society, whereas the authentic Pauline letters often take for granted that women are active in both spheres.¹²

2.2 Style and the Origins of the Debate

In addition to the four issues listed above, style is also a key issue in discussions of the Pauline and Pastoral Letters. In fact, observed stylistic differences between both corpora were a major factor that set the earliest debates in motion. Friedrich Schleiermacher is among the earliest critics who disputed the authenticity of a Pauline letter. In 1807, Schleiermacher penned a public letter to J. Chr. Gass, wherein he questions the authorship of 1 Timothy to a large degree on linguistic grounds, in particular the large number of *hapax legomena* found in that letter.¹³ Johann Gottfried Eichhorn expanded Schleiermacher’s argument by claiming that all three Pastorals are inauthentic, a finding he says is supported by the presence of peculiar words and phrases found nowhere else in the Pauline corpus.¹⁴ Shortly thereafter, Wilhelm M.L. De Wette argued against the authenticity of the Pastorals, claiming that “one familiar with Paul’s style of writing will find much aside from the differences in vocabulary...that betrays an imitator, who writes not from the living fullness of the Apostle’s mind.”¹⁵ Toward the end of the nineteenth century, Heinrich Holzmann lodged what is seen as the watershed study in the history of criticism of the Pastorals.¹⁶ Holzmann advanced a set of primarily linguistic arguments and noted, for example, that the Pastorals lack 25 particles typical of Paul’s style and that they contain a suspicious abundance of compound words and idiosyncratic phrases found nowhere else in the Pauline corpus.¹⁷ From the inception of the debate, scholars have discussed the issue of the authorship of the Pastorals in no small measure from the perspective of style and have utilized letters deemed authentically Pauline as a touchstone to evaluate the authenticity of the disputed ones.¹⁸

⁹ For example, Dunn, *Unity and Diversity in the New Testament*, 341–45; Hertzner, “Was ist falsch an der ‘fälschlich so genannten Gnosis?’”

¹⁰ For a discussion of the theology of the Pastorals with a view to its *Sitz im Leben*, refer to Dibelius and Conzelmann, *The Pastoral Epistles*, 8–10.

¹¹ Porter, “Pauline Authorship and the Pastoral Epistles,” 112.

¹² For a historical interpretation of the depiction of women in 1 Timothy, refer to Hylen, *A Modest Apostle*, 43–70.

¹³ Schleiermacher, *Über den sogenannten Ersten Brief des Paulus an den Timotheus*, 28–89. Schleiermacher argued that many of the unique terms in 1 Timothy derive from 2 Timothy and Titus.

¹⁴ Eichhorn, *Einleitung in das Neue Testament*, 315–7.

¹⁵ De Wette, *An Historico-Critical Introduction to the Canonical Books of the New Testament*, 301.

¹⁶ Porter, *The Pastoral Epistles*, 15; van Nes, *Pauline Language and the Pastoral Epistles*, 18.

¹⁷ Holzmann, *Die Pastoralbriefe, kritisch und exegetisch behandelt*, 84–118.

¹⁸ Since the inception of the discussion in German circles, there have been countless contributions, far too many to list here. Refer to Ehrman, *Forgery and Counterforgery*, 193–205, for an example of a recent treatment of differences in language between the

2.3 Statistical Approaches to the Style(s) of Paul and the Pastorals

The earliest debates on the authenticity of the Pastoral Letters eventually gave rise to statistical analyses of their style(s). These studies, like the ones strictly in the qualitative tradition, also utilize a selection of texts from the Pauline corpus as a standard of measurement against which the Pastorals can be compared. The first New Testament scholar to contribute a statistical analysis was Percival Neale Harrison, who found that “In vocabulary, grammar, and style...the Pastorals show a marked divergence from the other epistles bearing the name of Paul.”¹⁹ Harrison finds that the Pauline corpus consists of “two vocabularies, Pauline and Pastoral,” the latter of which is closer to the Apostolic Fathers than other New Testament authors.²⁰ Nevertheless, Harrison also finds that the Pastoral Letters contain a number of “Pauline fragments,” which he defines as excerpts of text originally written by Paul. Harrison arrives at his fragment hypothesis first by finding great stylistic variation between the Pastorals and the other ten letters in the Pauline corpus and second by noting that five sections within the Pastorals represent an exception. Harrison concludes that these five sections are authentic Pauline fragments because they are stylistically similar to the texts that he uses as a standard of measurement.²¹ In the wake of Harrison’s study, many other scholars deployed statistical approaches to the Pauline and Pastoral Letters, and the operative premise that Paul’s letters can be used as a standard of measurement has manifested itself in three main ways.

First, many computational studies take as their point of departure alleged linguistic differences observed by scholars in the qualitative tradition. For example, since the qualitative tradition has observed that the Pastorals contain a suspiciously high number of *hapax legomena* in comparison to the other Paulines, a range of statistical studies has emerged to investigate this issue. Kenneth Grayston and Gustav Herdan analyzed the statistical significance of unique words in the Pastorals by measuring what they term a *C*-value, which calculates the frequency of unique words in a corpus in terms of percentage. Grayston and Herdan found that the *C*-value of the Pastorals is 46%, an unusually high result when compared to the *C*-values of each of the individual Pauline letters, all of which are between 30 and 34%.²² Another alleged stylistic difference observed by scholars in the qualitative tradition is that the Pastorals have a rich vocabulary in comparison to the authentic Paulines. Some scholars, such as Matthew Brook O’Donnell,²³ Eta Linnemann,²⁴ and Gerard Ledger,²⁵ investigated this issue by computing Type-Token Ratios in the Pauline and Pastoral Letters. Jerro van Nes, in what is perhaps the most comprehensive study of the language of the Pauline and Pastoral Letters, uses linear regression analysis to re-examine the conclusions of previous scholars that the Pastorals are stylistically dissimilar from the authentic Paulines.²⁶ Van Nes considers a range of lexical and grammatical features that have been viewed throughout the history of criticism as evidence that the Pastorals were not written by Paul and concludes that most linguistic differences do not rise to the level of statistical significance. In our opinion, studies that take previously observed stylistic differences as their points of departure are useful in that they either verify or challenge conclusions reached in previous scholarship. We also contend, however, that basing a statistical analysis solely on previously observed differences imposes severe limitations on the types of stylistic features to be considered. Instead of taking stylistic peculiarities in the Pastorals as our starting point, we propose to begin with the “authentic” corpus to determine the mathematical distribution of

Pauline and Pastoral Letters with a view to authorship. For a literature review that discusses linguistic issues, refer to Schenk, “Die Briefe und Timotheus I und Timotheus II und an Titus (Pastoralbriefe) in der neueren Forschung (1945–1985),” 3404–38.

¹⁹ Harrison, *The Problem of the Pastoral Epistles*, 45.

²⁰ *Ibid.*, 68.

²¹ *Ibid.*, 96.

²² Grayston and Herdan, “The Authorship of the Pastorals in Light of Statistical Linguistics,” 1–15; for other computational studies that deal with *hapax legomena*, refer to Robinson, “Grayston and Herdan’s ‘C’ Formula and the Authorship of the Pastoral Epistles;” and Linnemann, “Echtheitsfragen und Vokabelstatistik,” 108; Morgenhalter, *Statistik des neutestamentlichen Wortschatzes*, 67–157.

²³ O’Donnell, “Linguistic Fingerprints or Style by Numbers?” 250.

²⁴ Linnemann, “Echtheitsfragen und Vokabelstatistik.”

²⁵ Ledger, “An Exploration of Differences in the Pauline Epistles Using Multivariate Statistical Analysis,” 93.

²⁶ van Nes, *Pauline Language and the Pastoral Epistles*.

features that are rarely considered in Pauline stylometry.²⁷ The advantage of this procedural order is that we would attain a firm standard of measurement: if we can determine whether stylistic features throughout Paul's letters follow a certain distribution, we can investigate how well the same features follow the same theoretical distribution in the Pastorals.

Second, especially in the mid-twentieth century, some studies make claims about the extent of the authentic Pauline corpus based solely on the results of statistical analysis. Andrew Queen Morton produced a number of statistical studies, one of which focused on sentence length and common words and concluded that only Romans, 1 and 2 Corinthians, Galatians, and possibly Philemon were written by Paul. Regarding the other letters, Morton finds that "the odds against Pauline authorship are astronomical."²⁸ Elsewhere Morton claims that the part of speech of the last word in a Greek sentence is the most critical factor in authorship attribution. Accordingly, Morton finds that only Galatians, 1 Corinthians, and parts of Romans are Pauline.²⁹ In both of his studies, Morton selects features based on what he thinks is most determinative of authorship and rejects the authenticity of any letter that emerges as an outlier. Anthony Kenny takes a similar approach in a stylometric study of the entire New Testament. Kenny's method of determining the authenticity of any letter of disputed Pauline authorship is "to start with the corpus of Pauline writings handed on by tradition, and ask whether within that corpus there is any Epistle, or group of Epistles, which is marked out as different from the body as a whole."³⁰ Based on 99 linguistic features,³¹ Kenny provides "a rough measure of how well each Epistle is at home in the Pauline corpus," which he lists in descending order: Romans, Philippians, 2 Timothy, 2 Corinthians, Galatians, 2 Thessalonians, 1 Thessalonians, Colossians, Ephesians, 1 Timothy, Philemon, 1 Corinthians, and Titus.³² Based on these results, Kenny finds that only the Pauline authorship of the letter of Titus is dubious: "I see no reason to reject the hypothesis that twelve of the Pauline Epistles are the work of a single, unusually versatile author."³³ We argue that, for both authors, there is an extreme over-reliance on statistical results to discriminate between authors. Methodologically, Morton and Kenny fail to consider that factors other than authorship, such as genre or topic, could account for stylistic variation within a given corpus.³⁴ On the opposite extreme, Kenneth Neumann concludes that stylistic variation within the literary output of a single author is common. As a result, Neumann is pessimistic that stylometry is an effective enterprise in forensic authorship attribution.³⁵ On the contrary, we argue that by fitting probability models to letters of undisputed Pauline authorship, we gain a statistical standard of comparison to be applied to the Pastorals and to be interpreted in light of scholarly core competencies.

Third, following the advent of multivariate statistical analysis in Pauline studies, scholars in recent decades have applied cluster analysis and related techniques to the Pauline corpus.³⁶ These methods enable the analyst to place the letters of Paul into groups based on stylistic similarity. The first such study was by D.L.

²⁷ For our process of features selection, see Section 3.

²⁸ Morton, *Literary Detection*, 182.

²⁹ Morton, "The New Stylometry."

³⁰ Kenny, *A Stylometric Study of the New Testament*, 95.

³¹ *Ibid.*, 123, lists all 99 features in an Appendix.

³² Kenny, *A Stylometric Study of the New Testament*, 98.

³³ *Ibid.*, 100.

³⁴ Libby, "The Pauline Canon Sung in a Linguistic Key," 127; Jockers, *Macroanalysis*, 63–104, shows how these and other factors are related in English novels.

³⁵ Neumann, *The Authenticity of the Pauline Epistles in Light of Stylostatistical Analysis*, 214.

³⁶ Among the related techniques, for example, is Pitts, "Style and Pseudonymity in Pauline Scholarship," who performs a register based analysis of the Pauline letters, dividing the 13-letter corpus into groups according to his understanding of their social backgrounds. Instead of providing new data on Paul's style, Pitts takes as his starting assumption that there is linguistic variation within the *Corpus Paulinum*, which he explains as a function of changes in social setting. Pitts finds that the greatest stylistic variation occurs between two groups of letters within the 13-letter corpus: 1) the Thessalonian correspondence, the *Hauptbriefe*, and the prison letters and 2) the Pastoral Letters. For Pitts, this result is to be expected, since the Pastorals exhibit the greatest degree of change in social circumstances from the other letters. Socio-linguistic studies such as the one by Pitts are important to the debate on the Pauline and Pastoral Letters, since these analyses produce explanatory models through which to interpret data on Paul's style. Since, however, our focus is on statistical models, we leave aside such studies for now, while acknowledging their relevance to the discussion.

Mealand, who analyzes 19 stylistic variables and concluded that the four major letters of Paul are stylistically distinct from the three minor ones.³⁷ James Libby applies correspondence analysis to the entire New Testament based on syntagmatic, paradigmatic, and semantic relationships within the Greek language. Libby concludes that 1 and 2 Thessalonians, Colossians and Ephesians, and 1 Timothy and Titus each form pairs, but that these clusters are more indicative of “epistolary subgenre” than authorship.³⁸ Jacques Savoy performs cluster analysis on term frequencies and found that the following letters within the Pauline corpus group together: Romans/Galatians/1 Corinthians/2 Corinthians, Colossians/Ephesians, 1 Thessalonians/2 Thessalonians, 1 Timothy/2 Timothy/Titus.³⁹ In our opinion, cluster analysis and related techniques are useful in so far as they respect the individuality of each letter and provide metrics of stylistic variation within the Pauline canon.⁴⁰ We contend, however, that even with multivariate methods, a major question remains unanswered: is it possible to determine how certain features behave across the entirety of the undisputed Pauline corpus, as opposed to assessing which features tend to occur only in letters of a certain length, in letters addressed to particular communities, or in letters that are close in theme or topic?

2.4 Synthesis

The question of how stylistic features behave in the undisputed letters generally goes unanswered in Pauline stylometry. We argue this is a major oversight in previous research, particularly since all inferential methods, including multivariate ones, rely upon an accurate understanding of the statistical distribution of stylistic features. However, instead of first determining the best-fitting distribution for a given set of stylistic features, researchers typically assume that their data follow a normal distribution without verifying that assumption. In New Testament stylometry, Mealand’s study is a notable exception, as he applies a univariate normality test to 25 variables and discards 6 of them.⁴¹ This result, however, problematizes whether the remaining variables are normally distributed: 24% of his variables (6 out of 25) were shown to be non-normal, almost five times higher than the 5% expected by the level of significance with which he works. The assumption of normality is required for Mealand’s discriminant analysis, but without confidence in this assumption, we cannot be confident in his conclusions.

By not adequately investigating the statistical behavior of stylistic features in the undisputed letters, scholars fail to assess how to utilize Paul’s letters as a standard of measurement. This oversight is potentially detrimental when applying inferential methods to letters of disputed authorship, because many statistical tests only produce valid results if the data follow a certain distribution. We intend, therefore, to improve upon previous studies by fitting our stylometric data to various probability distributions, the pay-off of which is that we can measure how certain features behave throughout Paul’s letters *before* turning to the Pastorals. Equipped with the knowledge of which probability distributions are the best fits to our data, we apply the same models to the Pastorals to evaluate whether there are statistically significant stylistic differences between these corpora.

To describe quantitatively Paul’s style, we consider the seven letters traditionally regarded as the undisputed Paulines: Romans, 1 and 2 Corinthians, Galatians, 1 Thessalonians, Philippians, and Philemon.⁴² To remain on secure ground when making claims about Paul’s style, we consider only these seven letters as “authentic,” since there are well-known problems with 2 Thessalonians, Colossians, and Ephesians. Previous studies have shown time and again that there is considerable stylistic variation even within the authentic

³⁷ Mealand, “The Extent of the Pauline Corpus.”

³⁸ Libby, “The Pauline Canon Sung in a Linguistic Key.”

³⁹ Savoy, “Authorship of Pauline Epistles Revisited.”

⁴⁰ This is arguably one of the major advantages of linear regression analysis as performed by van Nes, *Pauline Language and the Pastoral Epistles*. Van Nes plots each letter into vector space to visualize similarity or difference from the line of best fit.

⁴¹ Mealand, “The Extent of the Pauline Corpus,” 70.

⁴² For recent and authoritative expressions of this scholarly premise, refer to the essays in Wischmeyer, *Paul*.

letters (Section 2.3). We contend, however, that it is still possible to make statistically relevant claims about Paul's style at the level of his seven-letter corpus. Indeed, this approach is consistent with one of the foundational premises of stylometry, namely, that each author has a unique stylistic fingerprint that will always be left behind when writing a text.

3 Feature Selection

3.1 The Prominence of Vocabulary and Syntax

In this section, we explain how we select stylistic features to be analyzed. Statistical studies of the Pauline and Pastoral Letters almost exclusively revolve around lexical and syntactical features such as *hapax legomena*. In this way, previous contributions to the debate align with the way stylometry is typically practiced. Traditional stylometry operates with the premise that an author's stylistic "fingerprint" is detectable through the rates at which he or she deploys quantifiable linguistic features such as common function words.⁴³ Even studies that take more complex features as the basic unit of analysis nevertheless treat style primarily as a lexical and syntactical phenomenon. Donald Hardy, for example, analyzes the style Flannery O'Connor's novels through so-called collocational analysis, which measures the statistical significance of word associations.⁴⁴ Statistical studies of the Pauline and Pastoral Letters, then, have in common with other stylometric analyses their focus on vocabulary and grammar.

In recent years, however, a group of prominent voices in the digital humanities have called for a new definition of style, one that satisfies scholars who utilize traditional literary methods as well as those who engage in computational analysis. J. Berenike Hermann, Karina van Dalen-Oskam, and Christof Schöch define style as: "a property of texts constituted by an ensemble of formal features which can be observed quantitatively or qualitatively."⁴⁵ In explicating their definition, the authors note that in computational studies, "most style markers have so far been relatively simple in nature. Among such features are frequencies and frequency distributions of characters, words, lemmata, word classes or syntactical structures, taken by themselves or in sequences (n-grams)."⁴⁶ As a way forward, the authors argue that "it seems both useful and possible to add the levels beyond the sentence among the style markers."⁴⁷ As examples, they mention "stylistically relevant phenomena such as metaphors or irony which, despite some recent advances, are currently analyzed in a qualitative rather than quantitative paradigm and remain a challenge to formal modeling, reliable automatic identification and computational assessment."⁴⁸ Additionally, these authors even suggest it would be valuable if elements of a text's macro-structure such as paragraph length or point of view were given more prominence in the quantitative analysis of style.⁴⁹ In comparison to traditional stylometry, this is a broader and more inclusive approach to style, which in our opinion has the potential to shed new and informative light on Pauline and the Pastoral Letters. Since vocabulary and syntax have been given more than enough attention in the debate, we take a broad approach to style and compare both corpora in terms of a fresh set of literary techniques.

In recent years – independently, as far as we can tell, from the authors discussed above – a few voices in New Testament studies have similarly called for the integration of non-linguistic features into stylometric analysis. Matthew Brook O'Donnell, for instance, criticizes computational stylometry in biblical studies

⁴³ Foundational studies that analyze function words include Zipf, *Human Behavior and the Principle of Least Effort*; Mosteller and Wallace, *Inference and Disputed Authorship*; Burrows, "Delta."

⁴⁴ Hardy, "Collocational Analysis as a Stylistic Discovery Procedure," 410–27.

⁴⁵ Hermann et al., "Revisiting Style, a Key Concept in Literary Studies," 44.

⁴⁶ Ibid., 45.

⁴⁷ Ibid., 46.

⁴⁸ Ibid.

⁴⁹ Ibid.

because it fails to consider important non-linguistic elements in a text.⁵⁰ James Libby argues that working solely with traditional stylistic criteria is no longer sufficient, because such an approach provides only a limited picture of an author's style and because contemporary digital tools make possible different types of analysis.⁵¹ Paul Robertson theorizes that literary criteria in Paul's letters are best classified according to "second-order criteria," which he defines as terms and taxonomies the analyst develops for descriptive purposes and which may not have been shared by the cultural era in question.⁵² Although Robertson is not the only scholar who is dubious of traditional stylometry, he is to our knowledge the only one (within New Testament studies, at least) who engages in the quantitative analysis of style outside the parameters of vocabulary and syntax.⁵³ The categories that Robertson uses to compute Paul's style are as follows.⁵⁴

Universal Claims or Assertions	Pathos
Appeals to Authority	Irony or Satire
Conversation	Hyperbole
Prosopopoiia/Ethopoia	Oppositions or Choices
Rhetorical Questions	Figurations of Groupness
Metaphors or Analogues	Plural Inclusive Addresses
Anecdotes or Examples	Second Person Addresses
Imperatives	First Person Reflection
Exhortation	Analysis of Questions or Objections
Caustic Injunctions	Systematic Argument

According to Robertson, the features listed above are particularly characteristic of ethical-philosophical texts in the ancient Mediterranean world such as Paul's letters or Epictetus' *Discourses*.⁵⁵ In this study, we adopt Robertson's approach to feature selection but make no claims that certain literary features are characteristic of certain types of texts. Rather, we develop second-order categories specifically with an eye toward the debate about the Pauline and Pastoral Letters. Accordingly, we followed three general principles in developing categories to describe stylistic features in these corpora:

- First, we looked for literary techniques that describe in a general sense how Paul deploys letter-writing as a communicative tool. As a parallel to how traditional stylometry considers frequency rates of common function words such as "and," "but," or "the," we consider the statistical distributions of common literary techniques found in Paul's letters and find his stylistic fingerprint in the best fitting models.
- Second, in keeping with our goal to most effectively utilize Paul's letters as a standard for measurement, we looked for literary techniques that recur throughout the "authentic" corpus, as opposed to literary techniques that we know occur only in one or two letters. During the process of feature selection, we relied upon general impressions of what is characteristic of Paul's writing as we developed categories to describe various literary techniques. In subsequent stages of the research process, we tested our impressions mathematically.
- Third, we developed categories with an eye toward what is quantifiable. We specify, for example, that we only consider occurrences of self-fashioning when Paul explicitly uses the first-person (Section 3.2). Self-fashioning need not be understood strictly in terms of explicit claims that an author makes about himself. Theoretically, one could argue that Paul engages in self-fashioning with every word that he writes. In the interest of quantification, however, we limit ourselves to empirical text-features that can be counted and frame our categories accordingly.⁵⁶

⁵⁰ O'Donnell, *Corpus Linguistics and the Greek of the New Testament*, 85–101.

⁵¹ Libby, "The Pauline Canon Sung in a Linguistic Key."

⁵² Robertson, *Paul's Letters and Contemporary Greco-Roman Literature*, 141–52. For a summary of his method, see Robertson, "Visualizing Data in the Quantitative Comparison of Ancient Texts."

⁵³ Roy and Robertson, "Applying Cosine Similarity to Paul's Letters," 88–117.

⁵⁴ For a full discussion of these features, refer to Robertson, *Paul's Letters and Contemporary Greco-Roman Literature*, 121–69.

⁵⁵ Robertson, *Paul's Letters and Contemporary Greco-Roman Literature*, 121–69.

⁵⁶ For this reason, we find certain features from Robertson's list problematic, such as *pathos* and *conversation*, both of which we have struggled to quantify.

3.2 Cataloging Stylistic Features

Following the three principles outlined above, we developed 18 categories to describe stylistic features in Paul's letters. These 18 features do not encompass all aspects of Paul's letter-writing but are representative of many of the major ways the apostle relates to his audiences from a distance. The features are listed below along with a brief description of each:

- *Declarations of fact*: The author lodges a claim with wide-ranging implications in a manner that is declarative as opposed to, say, analytical.
- *Appeals to authority*: The author references a person, divine being, or text to bolster his own claims.
- *Logical arguments*: The author seeks to persuade his readers of the validity of an idea, or expounds upon an idea, with evidence and/or reasoning.
- *Inferences*: The author explicitly draws a conclusion based on what has proceeded.
- *First-person self-fashioning*: The author builds *ethos* by making claims about himself on topics such as his authority, past experiences, future plans, or exemplarity.⁵⁷
- *First-person emphases*: The author underscores his own voice by reinforcing that he is the one giving commands, offering opinions, and so forth, often with an emphatic pronoun or mention of his name.
- *First-person figurations of groupness*:⁵⁸ The author constructs a depiction of a community, whether concrete or abstract, that includes at least himself and his audience and that includes a description of things such as the group's ethical values, place in society, or eschatological destiny.
- *Second-person claims about audience*: The author depicts his readers in a specific way, also in terms of things such as their ethical values, place in society, or eschatological destiny.
- *Second-person emphases*: The author engages his readers by explicitly mentioning them, most often with an emphatic pronoun or the vocative case, even if doing so otherwise interrupts his flow of thought.
- *Third-person contrasts*: The author speaks disparagingly of others to construe them as outsiders or opponents and his focus is on the implied or explicit opposites with himself and/or his own group.
- *General exhortations or advice*: The author directs readers, usually with the hortatory subjunctive or imperative, to a general pattern of behavior, be it to adopt a new way of life, to maintain their current way of life, or to improve upon current performance.⁵⁹
- *Specific commands*: The author issues a particular directive to readers, as opposed to a general pattern of living, most often in the imperative mood.
- *Expressions of affection or friendship*: The author articulates warm feelings for and/or reinforces a sense of partnership with his audience, sometimes by deploying common tropes from ancient letter-writing (such as longing to see his readers), or even dismissing the fact of his physical absence by asserting, for example, his presence with his readers in spirit or through the letter.⁶⁰
- *Expressions of confidence*: The author articulates a sense of certainty in his audience, usually to perform a particular task or to live in accordance with a certain ethical framework.
- *Second or third-person praise*: The author positively evaluates a person, group, or virtue such as love, often to uphold the person or group as a moral exemplum because of, among other things, their past actions or present characteristics.
- *Second-person blame*: The author negatively evaluates his readers for something they have done or are doing, usually to prompt them to correct their behavior and/or to maintain a positive relationship with them.

⁵⁷ The expression self-fashioning was introduced into academic discourse by Greenblatt, *Renaissance Self-Fashioning*. For its application to ancient letter-writing, see e.g. Müller et al., *Adressat und Adressant in antiken Briefen*; Rosenmeyer, *Ancient Epistolary Fictions*.

⁵⁸ For the use of this term in sociology, refer to Robertson, *Paul's Letters and Contemporary Greco-Roman Literature*, 136–37.

⁵⁹ For discussion of the various modes of exhortation and their connections to advice, refer to Stowers, *Letter Writing in Greco-Roman Antiquity*, 94–152.

⁶⁰ The classic studies on letters as substitutes for personal presence are Koskeniemi, *Studien zur Idee und Phraseologie des griechischen Briefes bis 400 n. Chr.*; Thraede, *Grundzüge griechisch-römischer Briefftopik*.

- *Satire*: The author exposes what he considers to be flaws in his audience or others with ridicule, irony, or humor, usually as a form of public criticism on moral and/or social matters.⁶¹
- *Expressions of wish or desire*: The author articulates his yearning for something specific, usually in the optative mood and/or in the form of a prayer.

In the process of feature selection, we conducted multiple close readings of the Pauline and Pastoral Letters. After we identified generalized ways to describe how Paul uses letter-writing as a communicative tool, we decided where in our judgment these features occur in Paul's letters and catalogued text references accordingly.⁶² For example, we find that in Paul's letter to the Romans *inferences* occur 25 times and *first-person self-fashioning* occurs 10 times. Therefore, we catalogue these features in this way:

Stylistic feature	Text reference in Romans
<i>Inferences</i>	1:24–25; 2:1; 4:22; 5:1; 5:18; 6:4; 6:12; 7:12; 7:21; 7:25; 8:1; 8:12; 9:18; 10:17; 11:5–6; 13:2; 13:5; 13:10; 14:3; 14:12; 14:16; 14:19; 15:7; 15:17; 15:22
<i>First-person self-fashioning</i>	1:1–5; 1:13–15; 1:16; 9:1–3; 11:1; 11:7–9; 11:13–14; 15:16–20; 15:22–29; 16:1

This etic approach to feature selection contrasts with the way feature selection is conducted in traditional stylometry, which typically involves the use of corpus linguistics software to collect data for text properties such as word frequencies or clusters. The major advantage of engaging in feature selection through close readings is that we draw upon the expertise of scholars of the ancient world to identify and describe elements that are fundamental to the apostle's style, even if these elements are not explicitly acknowledged by Paul or by contemporary literary theorists.⁶³ In many of the features listed above, we describe stylistically relevant features of Paul's letters using grammatical, rhetorical, or theoretical concepts already in use in biblical studies and related fields.

In this connection, some of our features, such as *appeals to authority* and *first-person figurations of groupness*, are similar or identical to the ones Robertson uses in his work. In other cases, we think we have improved upon Robertson's categories. For example, Robertson includes as *first-person reflection* most instances where the apostle speaks about himself, which encompasses nearly every occurrence of the first-person singular.⁶⁴ By contrast, we distinguish between direct claims Paul makes about himself (Rom 11:13: “**I am** the apostle to the Gentiles”) from those instances when he highlights that he is the one speaking (Phil 4:3: “Yes, **I ask you also**, true companion, help these women [...]”). As a result, in our list of features, *first-person self-fashioning* is a separate literary technique from *first-person emphases*. Similarly, Robertson counts as *second person address* most instances where the author uses the second-person plural.⁶⁵ By contrast, we distinguish between direct claims Paul makes about his audience (1 Cor 6:11: “But **you were** washed, **you were** made holy, **you were** made righteous in the name of the Lord Jesus and by the Spirit of our God”) from those instances when he simply draws his readers into the discussion (1 Thess 4:13: “But we do not want you to be uninformed, **brothers**, about those who are asleep [...]”). As a result, in our list of features, *second-person claims about audience* is a separate literary technique from *second-person emphases*. In our approach to feature selection, stylistic features can be signaled by but are not defined by linguistic text-elements. For example, whereas in traditional stylometry the imperative mood might be considered a stylistic feature in its own right, in our study we judge whether Paul uses this mood to issue a *general exhortation or advice* or a *specific imperative*.

⁶¹ Although some ancient rhetorical theorists treat satire in a narrow sense of a specific literary genre, we treat satire in the broader sense as a mode of social discourse.

⁶² For our full catalogue of features, refer to Appendix.

⁶³ Robertson, *Paul's Letters and Contemporary Greco-Roman Literature*, 121.

⁶⁴ *Ibid.*, 138–9.

⁶⁵ *Ibid.*, 138.

Unsurprisingly, the process of cataloging features revealed many instances where two or more features co-occur in the same verse. In these cases, we simply register these verses as containing a plurality of literary techniques. In the figure below, we provide two examples of such overlap.

Text reference	Overlapping stylistic features	Text citation
Galatians 5:2	1. <i>First-person emphasis</i> 2. <i>Declaration of fact</i>	Behold, I, Paul, say to you that if you accept circumcision, Christ will be of no advantage to you.
1 Corinthians 10:14	1. <i>Inference</i> 2. <i>General exhortation or advice</i> 3. <i>Expression of affection or friendship</i>	Therefore, my beloved, flee from idolatry.

Also unsurprisingly, the process of cataloging confronted us with instances where we had to judge whether a particular section of text contains one long occurrence of a given feature or multiple occurrences of that same feature. Does Romans 12:9–13, for example, contain a single occurrence of *general exhortations or advice* or does it contain many successive occurrences of this feature? In contrast to Robertson, who tends to log such passages as a single occurrence, we consider a text like Romans 12:9–13 to contain multiple *general exhortations or advice*, since Paul issues a variety of separate directives. Similarly, we consider Galatians 4:6–9 to include two occurrences of *second-person claims about audience*. Although both occurrences have in common that the Galatians are no longer enslaved, we find that Paul makes two substantial and separate claims about his readers.

Text reference	Citation	Summary of contents
Gal 4:6–7	And because you are sons, God has sent the Spirit of his Son into our hearts, crying “Abba! Father!” So, you are no longer a slave, but a son, and if a son, then an heir through God.	You have been adopted as God’s sons through the Spirit.
Gal 4:8–9	Formerly, when you did not know God, you were enslaved to those that by nature are not gods. But now that you have come to know God, or rather be known by God, how can you turn back again to the weak and worthless elementary principles of the world, whose slaves you want to be once more?	You are returning to the state of slavery from which you have been redeemed by God.

3.3 Mathematical Considerations for Feature Selection

Our departure from traditional stylometry in terms of feature selection is a crucial step toward achieving our stated goal of assessing how to utilize Paul’s authentic letters as a standard of measurement. After cataloging all features according to the process described above, we divide the authentic Pauline corpus into 20-verse text-blocks and count the number of occurrences of each feature in each of the 74 blocks. In this way, our stylistic features are *discrete*, since they take on whole-number values. We may, for instance, observe 0, 1, or 2 occurrences of a stylistic feature in a text-block, but it is impossible to observe 1.5 occurrences. Discrete data are not well-modeled by the normal distribution, which is an example of a *continuous* random variable, meaning that a variable that is normally distributed is not constrained to be a whole number. In stylometry, the normal distribution is most often used to model *rates*, such as word frequencies.⁶⁶ Instead of

rates, our stylometric data involve *counts*, since our method involves organizing Paul's letters into text-blocks, which limits the occurrence of each literary technique to a small range of whole-number values. Since the normal distribution cannot model our stylometric data, we seek instead a model among the discrete probability distributions (Section 4).

With a dataset as small as Paul's authentic letters, a discrete probability model gives us a sharper understanding of the behavior of stylistic features than we would gain with a continuous one, since continuous probability models are more appropriate for larger datasets. Previously, D.L. Mealand sought to overcome the problem of the size of the New Testament letters by computing word counts in text-blocks of 1,000 words,⁶⁷ but this approach does not generate a good understanding of stylistic tendencies, since the small number of total text-blocks (35 blocks in 18 letters) results in small number of observations. By contrast, since our method renders 74 text-blocks across 7 letters, we have enough observations to fit the data to various distribution models. We argue, then, that considering epistolary modes of discourse enables us to make claims about Paul's style in his authentic letters, even though this corpus of letters is quite small.

3.4 Synthesis

We contend that our approach to feature selection offers two advantages that push the debate on the Pauline and Pastoral Letters into uncharted territory.⁶⁸ First, from the standpoint of how to define style, we argue that epistolary modes of discourse – such as whether Paul speaks of himself or about his readers, or whether he makes truth claims in a declarative or logical manner – are at least as stylistically relevant as word choice and other linguistic features. Since Pauline stylometry up until now has largely revolved around traditional features, we think a fresh perspective on style will enhance the discussion. Second, from a mathematical perspective, we argue that discrete probability models are critical to our pursuit of determining how to utilize Paul's letters as a standard of measurement, since we will see how literary techniques behave across the corpus of authentic letters. For these reasons, we modify an approach to stylometry theorized by Hermann and practiced by Robertson to advance the centuries-old debate on the Pauline and Pastoral Letters.

4 Paul's Letters as a Standard of Measurement

4.1 Probability Distributions

In this section, we fit probability distributions to stylistic features in Paul's letters. As indicated above, this marks a departure from most stylometric studies, which often use the normal distribution to model the rates of various features. Our approach, however, is in line with the approach taken by Mosteller and Wallace in their investigation of the authorship of the Federalist Papers.⁶⁹ We begin with a general introduction to probability distributions and their applications to ancient texts.

In stylometry, aspects of an author's style are observed and measured quantitatively. Nevertheless, these observations do not follow patterns that can be predicted with total accuracy in the way that, for example, the motion of a body under the influence of gravity can be predicted. Consequently, these patterns can only be described in terms of probability. In mathematical probability theory, a set of outcomes or observations, together with the probability of each outcome, makes up a *distribution*. A distribution is a mathematical description of a random variable. It can be used to describe the likelihood of different outcomes and make predictions based on that information. In our

⁶⁶ For a summary of how the normal distribution can be used to model frequency rates, refer to Williams, "Josephus, Stylometry, and Jewish Studies."

⁶⁷ Mealand, "The Extent of the Pauline Corpus."

⁶⁸ Robertson conducts stylometric analysis only with the seven undisputed letters and 2 Thessalonians.

⁶⁹ Mosteller and Wallace, *Inference in an Authorship Problem*.

investigation of Paul's letters, the outcomes are the number of times a given stylistic feature occurs in a text-block of 20 verses. To illustrate this concept, we provide a table which contains the distribution of *declarations of fact*.

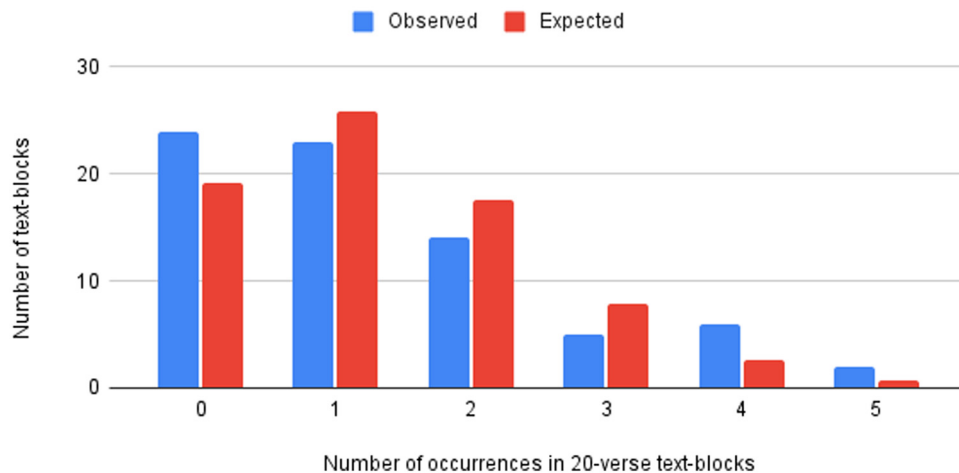
In the table below, the first column lists the number of occurrences of *declarations of fact* in a text-block, and the second column lists the number of text-blocks that contain that number of occurrences. For example, our data show that 24 text-blocks contain 0 occurrences of *declaration of fact*, while 23 text-blocks contain 1 occurrence. Finally, the third column records the expected number of text-blocks which contain the given number of occurrences of *declaration of fact* based on the chosen distribution. For this example, the expected counts in the third column are based on the Poisson distribution (Section 4.3).

Number of DOF	Observed	Expected
0	24	19.16
1	23	25.89
2	14	17.49
3	5	7.88
4	6	2.66
5	2	0.72

The strength of a probability model is assessed by its agreement between the observed values and the expected values. We visualize this distribution in the chart below, which shows general agreement between the observed and expected occurrences based on the Poisson model.

Declarations of Fact

Poisson model



When studying a stylometric variable, it is necessary to determine its distribution, since this information is useful for both stylometric *description* and *inference*. Regarding description, determining the distribution provides insight into how Paul uses a stylistic feature beyond what measures of the central tendency, such as the mean and median, and beyond what measures of the spread, such as the standard deviation and the variance, can tell us. While these statistics provide a basic summary of stylometric data, determining the distribution offers a more nuanced description of Paul's writing practices. Regarding inference, we compare the behavior of stylistic features in the authentic Paulines and the Pastorals and note quantitative differences. We ask whether the differences are significant – in other words, are the differences small enough that they could be explained in terms of the random variation found within the work of a single author? Or, conversely, are the observed differences so large that they are statistically unlikely to occur within the work of a single

author? In short, determining the distributions of stylistic features in Paul's letters helps us determine how these features behave in the authentic corpus, which in turn provides a standard of measurement in the quantitative comparison to the Pastorals.

In what follows, we describe three probability distributions and our method of fitting stylistic features to these models. Readers interested in the details of our statistical analysis are encouraged to read Section 4.2 on the binomial model, Section 4.3 on the Poisson model, and Section 4.4 on the negative binomial model. Alternatively, readers may safely skip to Section 4.5, where we display and discuss our results.

4.2 Assessing the Binomial Model

One of the simplest discrete probability models is the binomial distribution. This distribution models situations consisting of repeated trials, in which each trial results in exactly two outcomes, labeled “success” and “failure,” with a fixed probability of “success” and “failure.” The binomial distribution has the further requirement that all trials are independent. There is much to recommend this model, since the binomial setting has much in common with the occurrences of our stylistic features in Paul's letters. For example, each verse may be considered a trial, which would be a “success” if a given feature occurs and a “failure” if it does not.

The binomial distribution is determined by two parameters: n , which represents the number of trials, and p , which represents the probability of “success” in a trial. Since we organize Paul's letters into text-blocks of 20 verses, the value of the first parameter is $n = 20$. To determine the second parameter, we use the method of Maximum Likelihood Estimation, which states that the best-fitting binomial distribution arises when $p = \bar{x}/n$, the average value of the observations divided by the number of trials.

Using this method, we found the best-fitting binomial distribution for each stylistic feature. We assessed how well the binomial distribution models our observed data with χ^2 -Goodness-of-Fit tests. To conduct the χ^2 -Goodness-of-Fit test, we combined all text-blocks beyond a certain cut-off into a single category to meet the requirement that at most one category would have an expected value below 5.⁷⁰ A high p -value indicates that the binomial model is consistent with the observed data, while a low p -value suggests the binomial model is not consistent with the observed data. As the benchmark for statistical significance, we use the standard threshold of 5%, meaning we reject the binomial model for those features with a p -value below 5%. We record the p -value for each Goodness-of-Fit test in Section 4.5.

We see that the binomial distribution serves as a good model for some stylistic features, while for others it is rather poor. As a result, we consider in the following sections two additional discrete probability distributions, the Poisson distribution and the negative binomial distribution. The Poisson distribution improves upon the binomial in those cases when “successes” are rare, and the negative binomial improves upon Poisson when the data are “over-dispersed,” which occurs when large observations are common.

4.3 Assessing the Poisson Model

We begin by outlining some fundamental aspects of Poisson distributions, which can be found in many introductory textbooks on probability.⁷¹ A Poisson distribution describes the number of “successes” that occur in a fixed interval, provided that:

1. The probability of two or more successes in any sufficiently small interval is 0.
2. The probability of success is the same for any two intervals of equal length.

⁷⁰ Sullivan, *Statistics*, 589–91.

⁷¹ Refer to, for example, Sullivan, *Statistics*, 347–50.

3. The number of successes in any interval is independent of the number of successes in any other non-overlapping interval.

We claim that many stylistic features in Paul's letters meet these conditions:

1. If we take a small enough section of the text, there can be at most a single instance of any given stylistic feature.
2. As long as two text-blocks are of the same length, in many cases we have no reason to suspect that one text-block will contain more instances of a given feature than another text-block. There are some exceptions to this, which will occupy our attention when we turn to the negative binomial distribution (Section 4.4). For now, we simply note that this requirement will be true in many cases.
3. In most cases, we have no reason to suspect that the occurrence of one stylistic feature in one text-block should influence its occurrence in another block. Once again, there are exceptions, and these will become an important distinction between the Poisson and negative binomial models (Section 4.4).

A Poisson distribution is determined by a single parameter, λ , which represents both the mean and variance of the distribution. In our context, λ represents the average number of occurrences of a given stylistic feature in a text-block. To find the best-fitting Poisson distribution for a given set of observations, we determine the value of λ using the method of Maximum Likelihood Estimation. In the case of the Poisson distribution, this method states that the best-fitting Poisson distribution arises when λ takes the value of the mean of the data set.

To assess the strength of the Poisson model, we calculate the expected frequencies based on the value of λ . For any integer k , the probability of having exactly k occurrences of a given feature in a text-block is as follows:

$$P(k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Multiplying this probability by 74, the total number of 20-verse text-blocks in Paul's letters, yields the expected number of text-blocks with exactly k occurrences of the given feature. Having obtained the observed and expected frequencies of each feature, we conduct χ^2 -Goodness-of-Fit tests. The results for each feature are recorded in Section 4.5.

In the table in Section 4.5, note that two stylistic features, *appeals to authority* and *satire*, have two sets of χ^2 -statistics and p -values because we analyzed the data in two different ways. For *appeals to authority*, one text-block stands out as an outlier. The text-block consisting of Romans 10:5–11:3 contains 12 instances of this feature while most text-blocks contain 0–2 occurrences. For this reason, we employed the method of Maximum Likelihood Estimation twice, once using the value of λ obtained by using all text-blocks and another time omitting the outlier. The first value of λ yielded the p -value of $p = 26.4\%$, while the second value yielded $p = 66.04\%$. Both p -values are well above the cut-off from a 5% level of significance, meaning we may safely conclude that the Poisson distribution is a good model for Paul's use of *appeals to authority*, regardless of whether we retain or discard the outlier.

For *satire*, the first set of values results from organizing the text-blocks into three categories, those with no occurrences of *satire*, those with exactly 1 occurrence, and those with 2 or more occurrences. This is the standard way of organizing the data according to our method, since it results in exactly one category whose expected value is less than 5. The second χ^2 -statistic and p -value, recorded in parentheses, results from organizing the data into two categories, those with no occurrences of *satire* and those with 1 or more occurrences. This is not our standard way of organizing the data, but we have done so in this case because this alternative categorization yields a significantly higher p -value. Since such an increase in the p -value is caused by a small change in the way the data are categorized, we prefer the first p -value (0.04%), meaning we are not confident that the Poisson distribution is a good model for *satire* in Paul's letters.

4.4 Assessing the Negative Binomial Model

Next, we consider the negative binomial distribution, another discrete probability distribution that improves the binomial distribution when the outcome of “success” is rare. When we find that a stylistic feature is better modeled by the negative binomial distribution, we must ask what aspects of this stylistic feature make it differ from the Poisson distribution, since there are multiple settings that give rise to the negative binomial. Among these settings, we consider the presence of non-homogeneous and contagious populations to be the possibilities most relevant to Paul’s letters. Our treatment of these two settings, which we defer to Section 4.5, follows the landmark study of accident statistics by A.G. Arbous and J.E. Kerrich.⁷² Before we can consider why the negative binomial may arise, we must first determine how well the model works for our stylistic features.

We conduct χ^2 -Goodness of Fit Tests to assess the negative binomial distribution. As with the Poisson distribution, we first determine the parameters of the negative binomial distribution that best fit our observed data. The negative binomial distribution depends on two parameters, r and p , where r is a positive integer and $0 < p < 1$. For any integer $k \geq 0$, the probability of exactly k occurrences of a given stylistic feature in a text block is as follows:

$$P(k) = \binom{r+k-1}{k} (1-p)^k p^r.$$

For a given value of r , the method of Maximum Likelihood Estimation can be used to determine p , according to the equation:

$$p = \frac{r}{r + \bar{x}},$$

where \bar{x} is the mean of the observations. Unfortunately, the value of r cannot be determined by the method of Maximum Likelihood Estimation because it must take integer values, so we checked all values of r within a reasonable range to find the value that minimized the χ^2 -statistic. Once we determined the values of the parameters r and p , we calculated the expected frequencies and compared them with the observed frequencies in the same manner as the Poisson distribution. The results of the χ^2 -Goodness-of-Fit-Tests are recorded in the table contained in Section 4.5.

As in Section 4.3, *appeals to authority* and *satire* have two sets of χ^2 -statistics and p -values because we analyzed the data in two different ways, for the same reasons described above. Interestingly, the results are similar. Regardless of whether we include the outlier among *appeals to authority*, the resulting distribution yields a high p -value, which indicates that the negative binomial serves as a good model. For *satire*, the standard way of organizing the text-blocks into categories results in a small p -value, $p = 2.14\%$, below the conventional 5% cut-off. If an additional category is included, the p -value jumps to $p = 44.92\%$. As a result, we conclude that the negative binomial distribution is not a good model for Paul’s use of *satire*.

4.5 Results

Now that we have fitted our stylistic features to three distributions, we consider which distributions are most suitable for using Paul’s letters as a standard of measurement. We summarize our findings in the table below, which compares the p -values from our χ^2 -Goodness-of-Fit tests for each stylistic feature, based on the procedures described above.

⁷² Arbous and Kerrich, “Accident Statistics and the Concept of Accident Proneness,” 400–12.

Stylistic feature	Binomial (%)	Poisson (%)	NB (%)
<i>Declarations of fact</i>	1.80	6.76	86.66
<i>Appeals to authority</i>	1.94	26.4 (66.04)	71.47 (99.57)
<i>Logical arguments</i>	0.28	0.65	78.07
<i>Inferences</i>	59.77	50.72	50.66
<i>First-person self-fashioning</i>	0.03	0.26	60.27
<i>First-person self-emphases</i>	50.65	72.87	97.94
<i>First-person figurations of groupness</i>	93.62	91.01	90.99
<i>Second-person claims about audience</i>	34.79	39.65	99.82
<i>Second-person emphases</i>	1.84	4.31	38.40
<i>Third-person contrasts</i>	55.74	60.89	95.68
<i>General exhortations or advice</i>	0	3.3307×10^{-14}	3.56 (4.08)
<i>Specific imperatives</i>	0.88	1.90	94.30
<i>Expressions of affections or friendship</i>	0.13	0.51	30.79
<i>Expressions of confidence</i>	92.21	91.71	91.81
<i>Second or third-person praise</i>	0.54	0.70	10.58
<i>Second-person blame</i>	1.69×10^{-5}	1.047×10^{-4}	7.29
<i>Satire</i>	0.02	0.04 (36.71)	2.14 (44.92)
<i>Expressions of wish or desire</i>	9.31	11.09	62.82

Although the binomial distribution often provides a good fit, it is rarely a better model than the Poisson distribution. Moreover, in those few cases where it does offer some improvement, the improvement is very slight. For this reason, we will not pursue the binomial any further as a potential model. Next, we make conclusions as to which of the other two distributions, the Poisson or the negative binomial, is the stronger model for each literary feature. Our strategy is to generally prefer the Poisson distribution as long as the *p*-value for that model is above 5%. This preference comes from the fact that the Poisson distribution depends on a single parameter, unlike the negative binomial, which depends on two parameters, thereby making the Poisson distribution the simpler model. Thus, our rule-of-thumb is that the simpler model is the better one. Only when the *p*-value of the negative binomial model is approximately 10 times greater than the *p*-value of the Poisson model do we conclude that the negative binomial distribution serves as a better model. We summarize our results in the table below.

Poisson	Negative binomial	Neither
<i>Appeals to authority</i>	<i>Declarations of fact</i>	<i>General exhortations or advice</i>
<i>Inferences</i>	<i>Logical arguments</i>	<i>Satire</i>
<i>First-person self-emphases</i>	<i>First-person self-fashioning</i>	
<i>First-person figurations of groupness</i>	<i>Second-person emphases</i>	
<i>Second-person claims about audience</i>	<i>Specific imperatives</i>	
<i>Third-person contrasts</i>	<i>Expressions of affection or friendship</i>	
<i>Expressions of confidence</i>	<i>Second-person or third-person praise</i>	
<i>Expressions of wish or desire</i>	<i>Second-person blame</i>	

Eight stylistic features cannot be modeled by Poisson distribution but are a good fit for the negative binomial. To describe what this means in terms of Paul's style, we consider two possible settings that could give rise to this distribution.⁷³

⁷³ The following discussion was inspired by an in-progress article by Robertson and McCauley, which also uses the Poisson and negative binomial distributions to analyze style in Paul's letters.

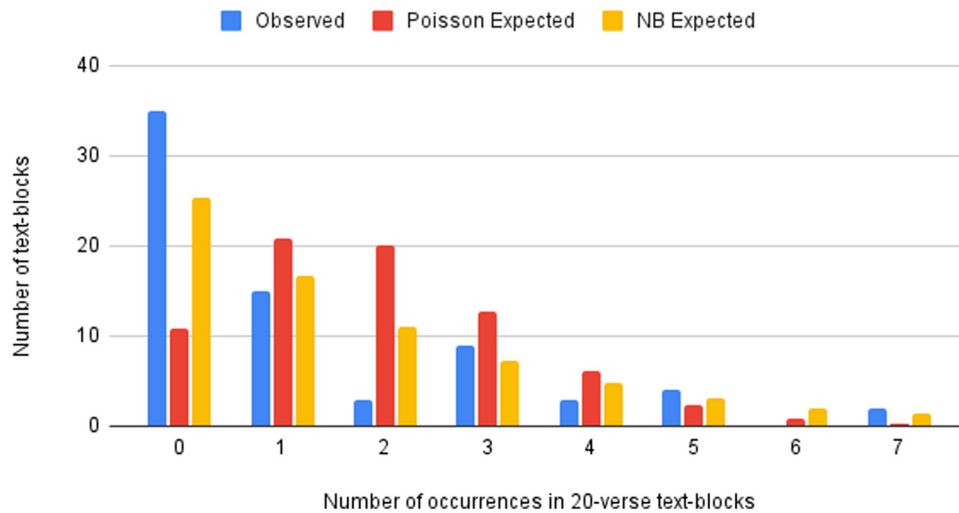
The first setting that creates a negative binomial distribution occurs when the population under study is not homogeneous. In other words, a population could consist of two or more sub-populations, each of which is homogeneous in the sense that it follows its own Poisson distribution. Under this scenario, the aggregate of these sub-populations, each with its own distinct Poisson distribution, would generate a negative binomial distribution. Theoretically, sub-populations within Paul's authentic letters could reflect the work of editors and interpolators, who supply their own styles governed by distinct Poisson distributions, so that the resulting mix would generate a negative binomial distribution. In our opinion, an alternative theory better explains why *expressions of affection or friendship* are a better fit for the negative binomial. This stylistic feature tends to occur somewhat formulaically in letter-openings as Paul offers greetings and gives thanks, although the letter to the Galatians is a notable exception. However, *expressions of affection or friendship* are also found throughout the main bodies of Paul's letters. Regarding Paul's style, this suggests that the letter-openings and main bodies of Paul's letters contain distinct Poisson distributions for *expressions of affection or friendship*, with the result that a negative binomial occurs in the aggregate.

The second setting that creates a negative binomial distribution is when the occurrence of one stylistic feature increases the probability that it will occur in a nearby text-block. This property has been called "contagious" because of the analogy with contagious disease, where the presence of disease in one individual increases the chances that it will occur in a nearby individual.⁷⁴ If a stylistic feature is contagious, its occurrences do not occur independently, which is a requirement for the Poisson distribution (Section 4.3). We think that seven of Paul's stylistic features fit the negative binomial model because they are contagious. Regarding *first-person self-fashioning*, for example, this means when Paul makes a claim about himself in a text-block, he is statistically more likely to make subsequent claims about himself in nearby text-blocks. Similarly, when Paul lodges a *logical argument*, he is more likely to put forth other *logical arguments* in the same context. Similarly, occurrences in one text-block of *declarations of fact*, *second-person emphases*, *second or third-person praise*, *second-person blame*, and *specific imperatives* in Paul's letters tend to heighten the probability of subsequent occurrences in nearby text-blocks. In other words, these features in Paul's letters have a cascading effect and, as a result, tend to cluster together in certain letter-parts. In some cases, contagious aspects of Paul's style can be observed at the macro-level of epistolary structure, such as Galatians 1–2 where *first-person self-fashioning* is a defining feature and Galatians 3–4 where *logical arguments* characterize the unit.

There remain two stylistic features that are not well-modeled by either distribution: *general exhortations or advice* and *satire*. Paul's use of *satire* seems to be a bad fit for these models because he uses this stylistic feature extremely rarely. Out of 74 text-blocks, 68 of them (roughly 92%) have no occurrences of *satire*. On the other hand, Paul uses *general exhortation or advice* very frequently, and many occurrences are highly concentrated in clusters, beyond what can be modeled by the Poisson or negative binomial. For example, the text-block corresponding to Romans 12:8–13:6 contains 23 occurrences of *general exhortations or advice* and the text-block corresponding to 1 Thessalonians 5:4–23 contains 13 occurrences. Aside from these clusters, the distribution of *general exhortations or advice* in Paul's letters significantly departs from the shapes of both theoretical Poisson and negative binomial distributions, in no small measure because the frequencies do not decrease steadily. To illustrate this point, we display a sample of the distribution of *general exhortations or advice* in the chart below.

⁷⁴ Although we use the term "contagious," it should be noted that this property describes successes that influence a greater or lesser number of successes in nearby intervals.

General Exhortations or Advice



4.6 Synthesis

In this section, we fit stylistic features in Paul's letters to various probability models to determine how to most accurately use the authentic letters as a standard of measurement. Most features are well-modeled either by the Poisson distribution or the negative binomial distribution, two distributions that improve upon the binomial when successes are rare. This result indicates that most stylistic features in Paul's letters occur only occasionally and that no feature occurs throughout his letters at a significantly higher rate than the others. Our finding may come as a surprise, for instance, to those who label Galatians as an "angry" letter or to those who see Philippians as a letter of friendship.⁷⁵ Statistically speaking, however, *second-person blame* and *expressions of affection or friendship*, like all of the stylistic features in our analysis, occur in a generally balanced manner throughout the authentic letters. This finding indicates that Paul in his letter-writing varied his style, consistently alternating between various epistolary modes of discourse, with the result that no stylistic feature dominates.

5 Application to the Authorship of the Pastoral Letters

5.1 Significance Tests

In this section, we compare the Pastoral Letters to the authentic Paulines in terms of the Poisson and negative binomial distributions that we observed in the previous section. We treat each of the Pastorals as individual letters instead of a collection because it is debated whether the Pastorals are the work of a single author or multiple authors. After controlling the lengths of each of the Pastorals, we conduct significance tests at the 5% benchmark to compare the observed number of occurrences of each feature in the Pastorals with the number we would expect based on the best-fitting distribution determined by the authentic Paulines. Since the Poisson

⁷⁵ For example, Stowers, "Friends and Enemies in the Politics of Heaven."

distribution is simpler than the negative binomial and since the negative binomial is the more conservative of the two, we only conduct negative binomial significance tests if there are statistically significant differences between Paul and the Pastorals based on the Poisson model. In running significance tests, our null hypothesis is that the Pastorals are stylistically similar to the authentic Paulines, in the sense that the stylistic features follow the same distributions as determined in Section 4.

In most cases, we fail to reject the null hypothesis that the Pastorals are stylistically similar to the authentic Paulines. This indicates that, for the most part, stylistic variation between the Pauline and Pastoral Letters is not statistically significant. This result challenges previous computational scholarship and the prevailing opinion in the qualitative tradition, which generally finds that the Pastoral Letters are stylistically distinct from the authentic Paulines (Section 2.3). However, two features in the Pastorals, namely *third-person contrasts* and *declarations of fact*, have p -values below 5%. In what follows, we explore the differences in the behavior of these features based on the Poisson distribution of *third-person contrasts* and both the Poisson and negative binomial distributions of *declarations of fact*. Although we determined above that *declarations of fact* are a better fit to the negative binomial (Section 4.5), we consider both distributions to demonstrate the consequences for statistical inference that arise based on one's selection of probability distribution.

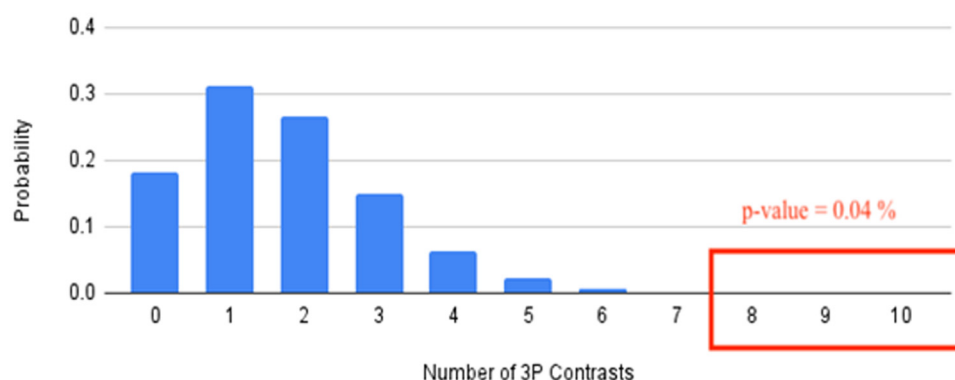
5.2 Poisson

5.2.1 Third-Person Contrasts

In the authentic Pauline letters, *third-person contrasts* occur at a mean rate of $\lambda = 0.30$ occurrences per text-block. Based on this information, we would expect Paul to deploy 1.70 occurrences of this stylistic feature in a text as long as 1 Timothy. By contrast, we observe 8 occurrences of this feature in 1 Timothy. We compute a p -value of 0.04%, meaning there is a 0.04% chance that the variation is due to random chance alone. Similarly, we would expect Paul to deploy 1.25 occurrences of *third-person contrasts* in a text as long as 2 Timothy. By contrast, we observe 7 occurrences of this feature in 2 Timothy. We compute a p -value of 0.03%, meaning there is a 0.03% chance that the variation is due to random chance alone. The single occurrence of *third-person contrasts* in the letter to Titus does not significantly differ from the 0.69 occurrences expected by the Poisson distribution we observed in Paul's authentic letters. We visualize these results in the charts below, placing in a box the portion of the right-tail that contributes to the p -value.

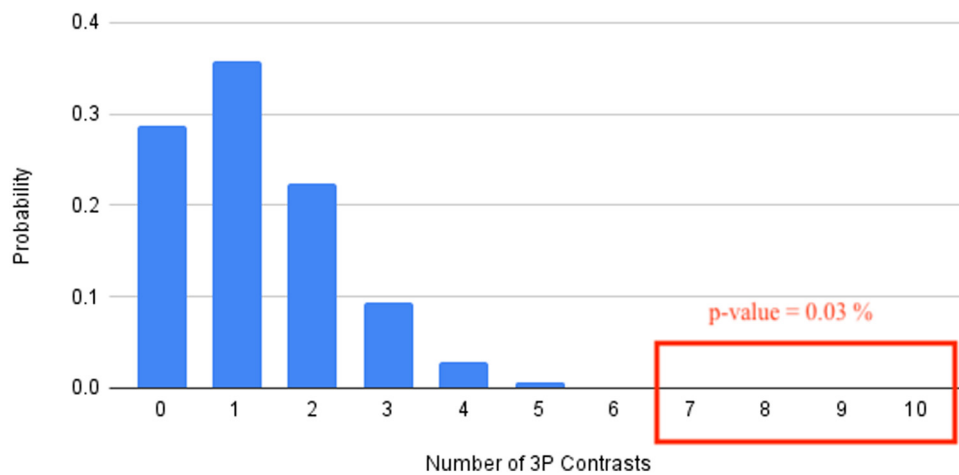
Expected 3P Contrasts in 1 Tim

Poisson distribution with $\lambda = 1.70$



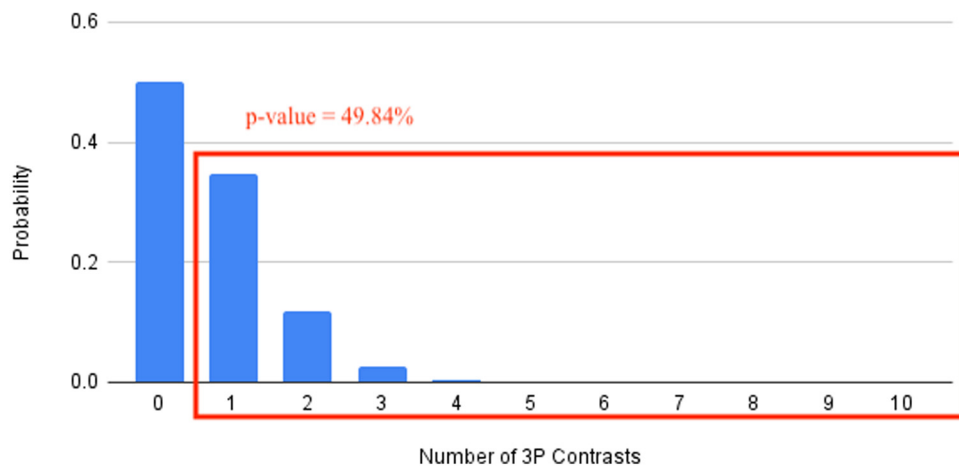
Expected 3P Contrasts in 2 Tim

Poisson distribution with $\lambda = 1.25$



Expected 3P Contrasts in Titus

Poisson distribution with $\lambda = 0.69$

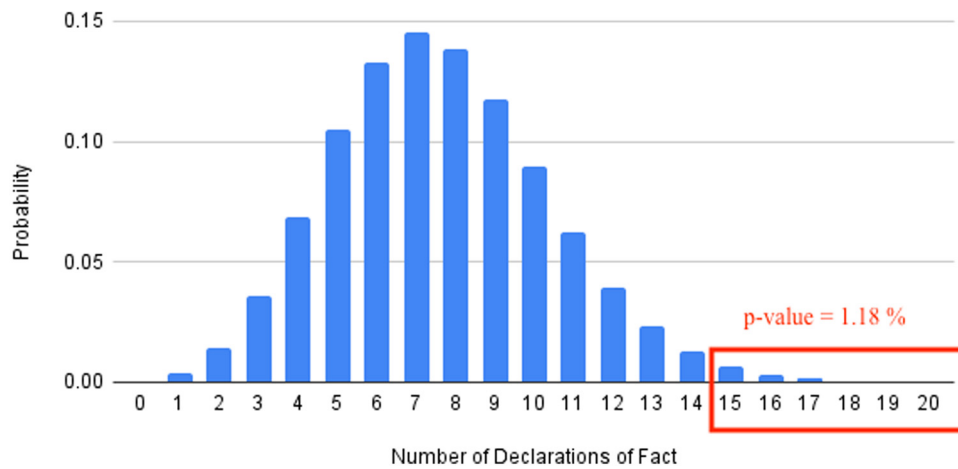


5.2.2 Declarations of Fact

In the authentic Pauline letters, *declarations of fact* occur at a mean rate of $\lambda = 1.35$ occurrences per text-block. Based on this information, we would expect Paul to deploy 7.63 occurrences of this stylistic feature in a text as long as 1 Timothy. By contrast, we observe 15 occurrences of this feature in 1 Timothy. We compute a p -value of 1.18%, meaning there is a 1.18% chance that the variation is due to random chance alone. Occurrences of *declarations of fact* in 2 Timothy or Titus do not significantly differ from the Poisson distribution we observed in Paul's authentic letters. We visualize these results in the charts below.

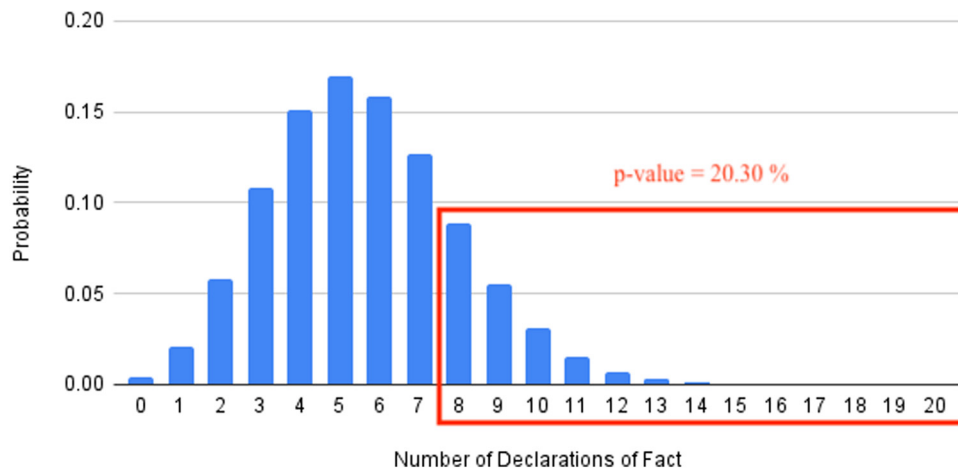
Expected Declarations of Fact in 1 Tim

Poisson distribution with $\lambda = 7.63$



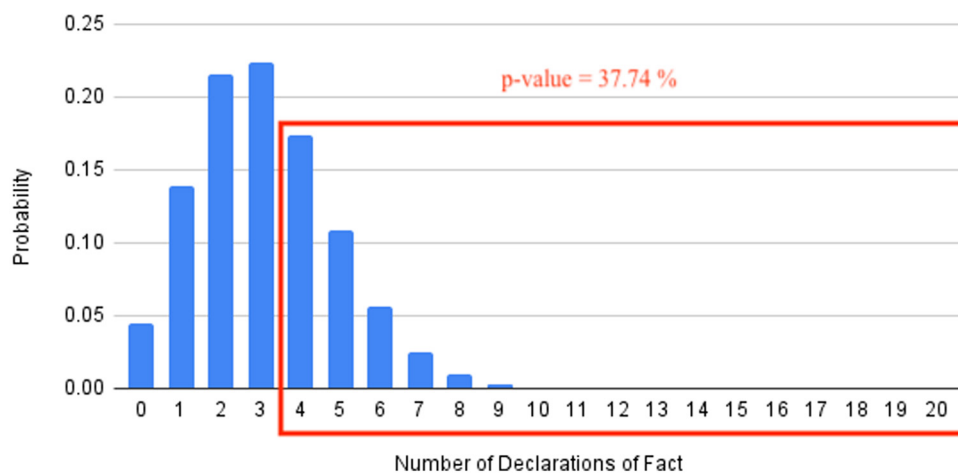
Expected Declarations of Fact in 2 Tim

Poisson distribution with $\lambda = 5.60$



Expected Declarations of Fact in Titus

Poisson distribution with $\lambda = 3.11$



5.3 Negative Binomial

In Section 4.5, we found that the negative binomial distribution is the best-fitting model for *declarations of fact* in Paul's letters. For this reason, we compute p -values for the Pastoral Letters using the negative binomial distribution. To highlight the importance of selecting the right distribution when making statistical inference, we compare the results with those obtained with the Poisson distribution (Section 5.2.3).

The procedure for conducting a significance test with a negative binomial distribution involves a few considerations not relevant to the Poisson distribution. Recall that the Poisson distribution depends on a single parameter, λ , which represents the average number of occurrences of a stylistic feature within a 20-verse text block. Therefore, when comparing the Pastorals with the authentic Pauline letters, we scaled the average rate for a 20-verse text-block to determine the expected number of occurrences of *declaration of fact* based on the length of each Pastoral Letter. This average value allowed us to determine λ and consequently the Poisson distribution for each Pastoral letter.

The negative binomial distribution relies on two parameters, r and p , neither of which can be determined directly from the average number of occurrences of a stylistic feature. For this reason, the average rate of occurrence cannot be scaled to determine the two parameters, meaning we must seek another way to find r and p . Our method is based on the following property of negative binomial random variables: suppose X_1 and X_2 are two random variables that follow a negative binomial distribution with two values of the first parameter, r_1 and r_2 , and the same value of the second parameter, p . Under this scenario, their sum, $X_1 + X_2$, follows a negative binomial distribution, in which the value of the first parameter is $r_1 + r_2$ and the second parameter is p . This result can be extended to the sum of any number of negative binomial random variables.⁷⁶

Applying this property to the Pastoral Letters allows us to extend the negative binomial distribution of features from Paul's authentic letters. To illustrate, we consider 1 Timothy, which contains 113 verses. Starting from the beginning of the letter, we divide 1 Timothy into 5 text-blocks of 20 verses, discarding the last 13 verses. If we consider the number of *declarations of fact* in a 20-verse text-block of the authentic Paulines to be a random variable X , then X follows a negative binomial distribution with $r = 2$ and $p = 0.597$. If we sum five negative binomial random variables, each with this same distribution, the result would be a negative binomial random variable with $r = 10$ and $p = 0.597$. If 1 Timothy is stylistically similar to Paul's authentic letters, the occurrences of *declarations of fact* would be governed by this negative binomial distribution, which we have determined to be the most accurate standard of measurement.

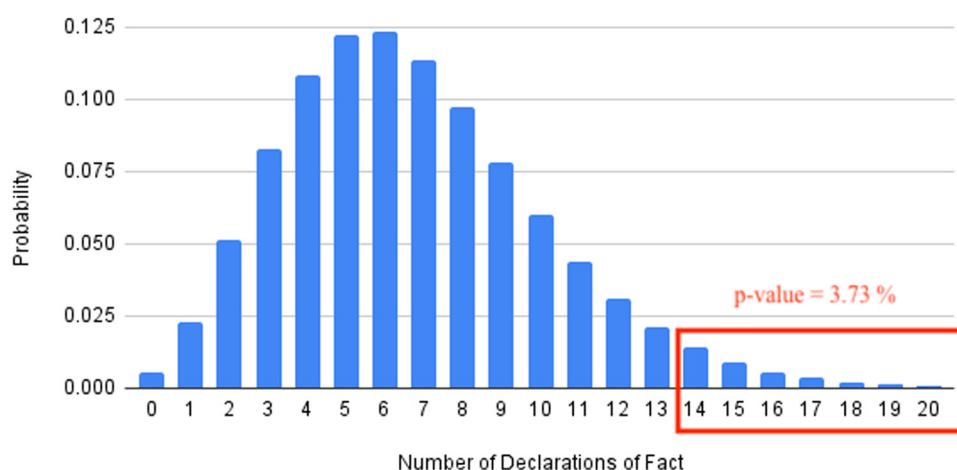
To test whether 1 Timothy deviates from Paul's style, we calculate the probability that we would observe as many as 14 instances of *declarations of fact* under the negative binomial distribution observed in Paul's letters.⁷⁷ This probability is the p -value, which we calculate as $p = 3.73\%$. Following this method, we have 4 text-blocks in 2 Timothy and a total of 8 occurrences of *declarations of fact*, and 2 text-blocks in Titus and a total of 4 occurrences. For 2 Timothy, we find p -value = 22.02%, and for Titus we find p -value = 29.56%. We visualize these results in the charts below.

⁷⁶ This is a standard property of negative binomial random variables, which can be found in many surveys of the negative binomial distribution. Refer to Girondot and Barry, "Computation of the Distribution of the Sum of Independent Negative Binomial Random Variables."

⁷⁷ We have observed 15 *declarations of fact* in our 5 text-blocks, but one occurrence is discarded, since it is found in the final 13 verses.

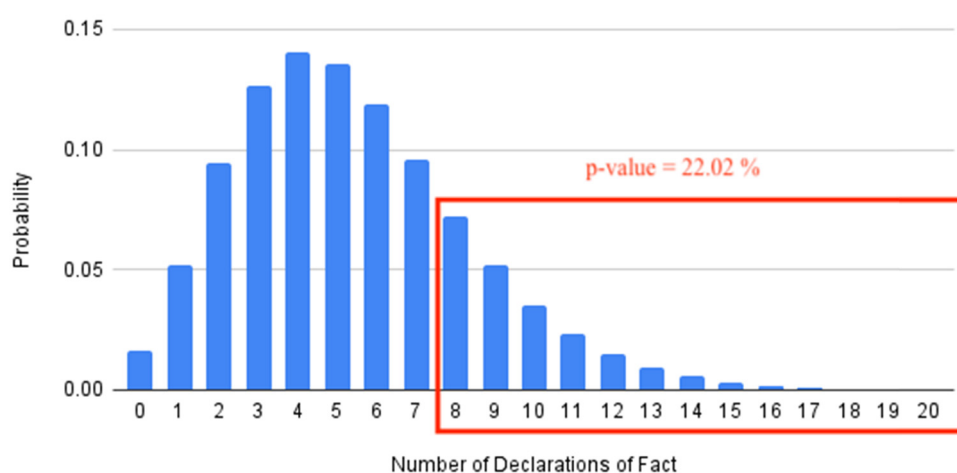
Expected Declarations of Fact in 1 Tim

Negative binomial with $r = 10$ and $p = 0.597$



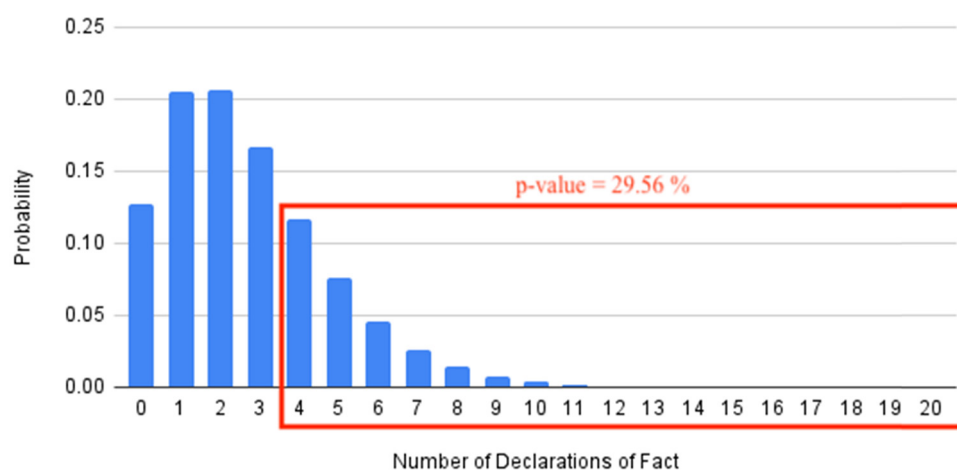
Expected Declarations of Fact in 2 Tim

Negative binomial with $r = 8$ and $p = 0.597$



Expected Declarations of Fact in Titus

Negative binomial with $r = 4$ and $p = 0.597$



On the surface, our results are similar to those produced by significance tests using the Poisson distribution, with 1 Timothy containing a sufficiently high number of *declarations of fact* to render the difference statistically significant, while 2 Timothy and Titus have large p -values, suggesting that *declarations of fact* in these letters do not depart from Paul's style as observed in the authentic letters. Although the conclusions for both the Poisson and negative binomial significance tests are similar, it is instructive to compare the p -values for 1 Timothy, both of which were below the 5% significance threshold. The p -value calculated under the negative binomial distribution is more than three times larger than the one found under the Poisson distribution, which reflects the fact that the negative binomial distribution has a larger right tail than the Poisson distribution, precisely the reason the negative binomial distribution is often considered an "over-dispersed" Poisson distribution. Thus, it is not unusual in a negative binomial distribution to find a large observation, such as the high number of *declarations of fact* we observe in 1 Timothy. These results show that researchers may reach vastly different conclusions depending on how they understand the underlying distribution of their data, and which tests they employ as a result. Our exercise reveals a broader lesson for Pauline stylometry, that determining how to use Paul's letters as a standard of measurement is an essential (and generally overlooked) first step in making statistical inferences.

6 Conclusion

In this study, we took as our point of departure two desiderata in the debate on Paul and the Pastorals: 1) the field lacks clarity on how to utilize Paul's letters as a standard of measurement and 2) the field over-analyzes the same types of stylistic features. To ameliorate these conditions, we cataloged the occurrences of 18 epistolary modes of discourse in the Pauline and Pastoral Letters. Once we had amassed a database of stylistic features, we fit our data to various probability models and found that 8 are a good fit to the Poisson distribution, 8 are a good fit to the negative binomial distribution, and 2 are not a good fit to either distribution. We have combined the core competencies of a New Testament scholar and a mathematician, who have worked together to advance the debate on the Pauline and Pastoral Letters. Our results are relevant to both statistical description and inference. The ways Paul relates to his audiences – whether declaring facts or making arguments, fashioning communal identity or disparaging outsiders, giving praise or issuing blame – can be quantified as features that exhibit certain statistical behaviors. By fitting these features to various probability models, we obtain an accurate standard of measurement by which we assess how the style(s) of the Pastorals compare to what we observed in the authentic Paulines. We conclude that stylistic differences between the Pauline and Pastoral Letters rise to the level of statistical significance only in a few cases, a finding that challenges many previous studies.

In this study, we have quantitatively described Paul's style. The question of which distribution is appropriate has largely been ignored in Pauline stylometry, with researchers often assuming, implicitly or explicitly, that the data under consideration follow a normal distribution. In our opinion, distribution fitting is an essential first step that must be performed before the analyst engages statistical inference. By determining which distribution provides the best model, based on theoretical hypotheses and verified through Goodness-of-Fit tests, we gain confidence in the results of the significance tests that we applied to the Pastorals, since we have first assessed how to use Paul's letters as a standard of measurement. As indicated above, our results dispute the findings of many previous studies, which generally find (or assume) significant differences in the styles of the Pauline and Pastoral Letters. However, our literature review has shown that two major computational studies in recent years – Jermo van Nes's monograph and James Libby's article – have concluded that stylistic differences between both corpora are not as significant as has previously been thought (Section 2.3). Based on our results, we agree with the findings of van Nes and Libby and argue that the differences in style between the Pauline and Pastoral Letters are not sufficient to discriminate between authors.

Crucially, our analysis differs from previous studies in the debate on the Pauline and Pastoral Letters because of the nature of the stylistic features we have analyzed. Mathematically, our feature selection is a step toward overcoming one of the most ubiquitous problems in Pauline stylometry: as samples of an author's writing habits, Paul's letters are small datasets. By considering features quantified by counts and modeled by

discrete probability distributions, we make 74 observations about the behavior of Paul's style and thereby gain more confidence in our findings than was possible in Mealand's landmark study, which could only consider 35 text-blocks across 18 New Testament letters (Section 3.3). To be sure, we have not overcome completely the issue of the length of Paul's letters, which will always be a small dataset regardless of how many troubleshooting maneuvers analysts may apply. Nevertheless, we contend that with our approach we gain more confidence than was previously possible in making inferences, which results in no small measure from selecting features that can be reasonably modeled in 20-verse text-blocks, thereby enabling us to observe Paul's writing habits across many intervals.

Our stylistic features also admit to the debate of an alternative approach to style. Although other scholars before us have theorized that stylometric analysis should not be limited to vocabulary and syntax, we execute an emerging approach and apply it to a long-standing authorship debate in New Testament studies. Style involves far more than traditional features such as term frequencies, lexical richness, sentence length, word positions, and syntax, all of which have been examined time and again in previous scholarship. Style can also be defined on the conceptual level to include the ways an author positions himself in relation to his readers. We have found that various epistolary modes of discourse occur throughout Paul's letters according to distributional patterns, which suggests that Paul is generally consistent in terms of how he relates to his readers. Paul's letters demonstrate great variety in terms of the situations he addresses, the tones he strikes, and the topics he discusses, all of which are considerations that add significance to our finding that Paul is consistent in how he deploys stylistic features.

Although we have shown there to be stylistic consistency across the authentic Paulines in terms of our 18 features, we acknowledge that future studies are required to assess the extent to which the distributions of these features are truly indicative of Paul's unique stylistic fingerprint. By design, the second-order categories we have formulated are quite general (Section 3.1), and we presume that each of these features recur in ancient letter-writing more generally. To evaluate further the extent to which the distributions of these features in Paul's authentic letters reveal the apostle's unique writing habits, these same features should be analyzed not only in other early Christian letters (e.g., 1 Peter, Ignatius, 1 Clement) but also in the letters of authors such as Cicero, Seneca, and Pliny. Such studies of epistolary modes of discourse in other authors may reveal overarching stylistic trends in ancient letter-writing or trends unique to, say, letters written under certain circumstances, letters with a certain tone, or letters by authors who share a common religious-philosophical outlook. Preparations for such investigations would be laborious since much more hand-coding would be required – and, for this reason, such a large-scale study is beyond the scope of this article – but such lines of inquiry have great potential to elucidate further the relationship between the Pauline and Pastoral Letters by showing the statistical behavior of epistolary modes of discourse in texts relevant to their literary milieu.

Even though the question of the distributions of our 18 features in other letters remains open, our study puts into perspective the role of style in debates on the Pauline and Pastoral Letters. As indicated above, two studies in recent years also challenge the narrative that there are major differences in the styles of both corpora. It appears, then, that whether style is understood in terms of vocabulary and syntax or in terms of epistolary modes of discourse, a scholarly trend is emerging in computational studies of the Pauline and Pastoral Letters that questions whether style is a sufficient criterion by which to discriminate between authors. In this connection, another avenue for future research in Pauline stylometry could be to combine lexical and syntactic features with the features identified in the article, or with other stylistically relevant elements also defined at the conceptual level. To be sure, the Pastorals pose many problems for Pauline authorship outside of the frame of style (Section 1), and we do not claim that our results show that Paul is the author of the Pastorals. However, since there are many other problems with the Pastorals, such as theological differences or historical settings, we find that arguments against Pauline authenticity based on style are comparatively weak.

Acknowledgments: We thank those who have read and provided constructive feedback to previous drafts: Paul Robertson (New Hampshire), Miriam DeCock (Dublin), and Christian Thrue Djurslev (Aarhus). We also thank the participants of both the Aarhus-Oslo-Münster NT meeting in Münster, Germany (Dec 2023) and the Workshop on Cognitive Approaches to Religion in Helsinki, Finland (May 2024) for productive discussions of earlier drafts.

Funding information: Research for this publication was supported by the Carlsberg Foundation in the Semper Ardens Accelerate project Computing Antiquity: Computational Research in Ancient Text Corpora.

Author contributions: The project was conceptualized by both authors. Erich Benjamin Pracht researched the history of the Paul and Pastorals debate, produced the database of stylistic features, and assisted in the statistical analysis. Thomas McCauley developed the mathematical aspects of the paper and implemented statistical procedures accordingly. Each author read and provided substantial input to all portions of the paper, including those outside of his core area of scholarly expertise.

Conflict of interest: The authors declare no conflict of interest.

Data availability statement: Our statistical procedures are freely available in spreadsheets at github.com/erichpracht.

References

- Arbous, Adrian Garth and John Edmund Kerrich. "Accident Statistics and the Concept of Accident Proneness." *Biometrics* 7 (1951), 340–432.
- Burrows, John. "Delta: A Measure of Stylistic Difference and a Guide to Likely Authorship." *Literary and Linguistic Computing* 17:3 (2002), 267–87.
- De Wette, Wilhelm M. L. *An Historico-Critical Introduction to the Canonical Books of the New Testament*, translated by Frederick Frothingham. Boston: Crosby, 1958.
- Dibelius, Martin and Hans Conzelmann. *The Pastoral Epistles: A Commentary on the Pastoral Epistles. Herm*, translated by Philip Buttolph and Adela Yarbro. Philadelphia: Fortress, 1972.
- Dunn, James D. G. *Unity and Diversity in the New Testament: An Inquiry Into the Character of Earliest Christianity*. Philadelphia: Westminster, 1977.
- Ehrman, Bart D. *Forgery and Counterforgery: The Use of Literary Deceit in Early Christian Polemics*. Oxford: Oxford University Press, 2012.
- Eichhorn, Johann Gottfried. *Einleitung in das Neue Testament*. Leipzig: Weidmannischen Buchhandlung, 1812.
- Girondot, Marc and Jon Barry. "Computation of the Distribution of the Sum of Independent Negative Binomial Random Variables." *Mathematical and Computational Applications* 28:63 (2023), 1–9.
- Grayston, Kenneth and Gustav Herdan. "The Authorship of the Pastorals in Light of Statistical Linguistics." *New Testament Studies* 6:1 (1959), 1–15.
- Greenblatt, Stephen. *Renaissance Self-Fashioning: From More to Shakespeare*. Chicago: University of Chicago Press, 1980.
- Hardy, Donald E. "Collocational Analysis as a Stylistic Discovery Procedure: The Case of Flannery O'Connor's *Eyes*." *Current Trends in Stylistics* 38:4 (2004), 410–27.
- Harrison, Percival Neale. *The Problem of the Pastoral Epistles*. London: Oxford University Press, 1921.
- Hermann, J. Berenike, Karina van Dalen-Oskam, and Christof Schöch. "Revisiting Style, a Key Concept in Literary Studies." *Journal for Literary Theory* 9:1 (2015), 25–52.
- Hertzer, Jens. "Was ist falsch an der 'fälschlich so genannten Gnosis'? Zur Paulusrezeption des Ersten Timotheusbriefes im Kontext seiner Gegnerpolemik." *Early Christianity* 5:1 (2014), 35–67.
- Holtz, Gottfried. *Die Pastoralbriefe*. THNT 13. Berlin: Evangelische Verlagsanstalt, 1965.
- Holzmann, Heinrich Julius. *Die Pastoralbriefe, kritisch und exegetisch behandelt*. Leipzig: Verlag von Wilhelm Engelmann, 1880.
- Hulen, Susan. *A Modest Apostle: Thecla and the History of Women in the Early Church*. Oxford: Oxford University Press, 2020.
- Jeremias, Joachim. *Die Briefe an Timotheus und Titus*. Göttingen: Vandenhoeck and Ruprecht, 1947.
- Jockers, Matthew L. *Macroanalysis: Digital Methods and Literary History*. Champaign: University of Illinois Press, 2013.
- Johnson, Luke Timothy. *The First and Second Letters to Timothy: A New Translation with Introduction and Commentary*. AB 35A. New Haven: Yale University Press, 2001.
- Kelly, John Norman Davidson. *A Commentary on the Pastoral Epistles: I Timothy, II Timothy, Titus*. New York: Harper and Row, 1963.
- Kenny, Anthony. *A Stylometric Study of the New Testament*. Oxford: Clarendon, 1986.
- Knight III, George W. *The Pastoral Epistles*. NIGTC. Grand Rapids: Eerdmans, 1992.
- Koehling, Heinrich. *Der Erste Briefe Pauli an Timeotheus: Auf's Neue Untersucht und Ausgelegt*. Berlin: Verlag von Hugo Rother, 1882.
- Koskeniemi, Heikki. *Studien zur Idee und Phraseologie des griechischen Briefes bis 400 n. Chr.* Helsinki: Suomalainen Tiedeakatemia, 1956.
- Kühl, Ernst. *Die Gemeindeordnung in den Pastoralbriefen*. Berlin: Wilhelm Hertz, 1885.
- Ledger, Gerard. "An Exploration of Differences in the Pauline Epistles Using Multivariate Statistical Analysis." *Literary and Linguistic Computing* 10:2 (1985), 85–97.

- Libby, James. "The Pauline Canon Sung in a Linguistic Key: Visualizing New Testament Text Proximity by Linguistic Structure, System, and Strata." *Biblical and Greek Linguistics* 5 (2016), 122–201.
- Linnemann, Eta. "Echtheitsfragen und Vokabelstatistik." *Jahrbuch für Evangelikale Theologie* 10 (1996), 99–109.
- Mealand, David L. "The Extent of the Pauline Corpus: A Multivariate Approach." *Journal for the Study of the New Testament* 18 (1995), 61–92.
- Morgenhalter, Robert. *Statistik des neutestamentlichen Wortschatzes*. Zürich: Gotthelf-Velag, 1958.
- Morton, Andrew Queen. "The New Stylometry: A One-Word Test of Authorship for Greek Writers." *The Classical Quarterly* 22:1 (1972), 89–102.
- Morton, Andrew Queen. *Literary Detection: How to Prove Authorship and Fraud in Literature and Documents*. New York: Scribner, 1978.
- Mosteller, Frederick and David Wallace. *Inference and Disputed Authorship: The Federalist*. Reading: Addison-Wesley, 1964.
- Mounce, William D. *Pastoral Epistles*. WBC 46. Waco: Word, 2000.
- Müller, Gernot Michael, Sabine Retsch, and Joanna Schenk, eds. *Adressat und Adressant in antiken Briefen: Rollenkonfigurationen und kommunikative Strategien in griechischer und römischer Epistolographie*. BA 382. Berlin: De Gruyter, 2020.
- Neumann, Kenneth. *The Authenticity of the Pauline Epistles in Light of Stylostatistical Analysis*. Atlanta: Scholars Press, 1990.
- O'Donnell, Matthew Brook. "Linguistic Fingerprints or Style by Numbers? The Use of Statistics in the Discussion of Authorship of New Testament Documents." In *Linguistics and the New Testament: Critical Junctures*, edited by Stanley E. Porter and D. A. Carson, 206–62. JSNTSS 168. Sheffield: Sheffield Academic Press, 1999.
- O'Donnell, Matthew Brook. *Corpus Linguistics and the Greek of the New Testament*. NTM 6. Sheffield: Sheffield Phoenix Press, 2005.
- Pitts, Andrew W. "Style and Pseudonymity in Pauline Scholarship: A Register Based Configuration." In *Paul and Pseudepigraphy*, edited by Stanley E. Porter and Gregory P. Fewster, 113–52. Leiden: Brill, 2013.
- Porter, Stanley E. "Pauline Authorship and the Pastoral Epistles: Implications for Canon." *Bulletin for Biblical Research* 5 (1995), 105–23.
- Porter, Stanley E. *The Pastoral Epistles: A Commentary on the Greek Text*. Grand Rapids: Baker, 2023.
- Prior, Michael. *Paul the Letter-Writer and the Second Letter to Timothy*. JSNTSS 23. Sheffield: Sheffield Academic Press, 1989.
- Robertson, Paul. "Visualizing Data in the Quantitative Comparison of Ancient Texts: A Study of Paul, Epictetus, and Philodemus." In *Ancient Manuscripts in Digital Culture: Visualization, Data Mining, Communication*, edited by David Hamidovic, Claire Clivaz, and Sarah Bowen Savant. DBS 3. Leiden: Brill, 2019.
- Robertson, Paul. *Digital Approaches to Paul's Letters: Studies and Applications in Statistics, Data Visualization, Topological Methods, and Higher Order Theories*. Leiden: Brill, 2026.
- Robertson, Paul. *Paul's Letters and Contemporary Greco-Roman Literature: Theorizing a New Taxonomy*. NTS 167. Leiden: Brill, 2016.
- Robinson, Thomas. "Grayston and Herdan's 'C' Formula and the Authorship of the Pastoral Epistles." *New Testament Studies* 30:2 (1984), 282–8.
- Rosenmeyer, Patricia A. *Ancient Epistolary Fictions: The Letter in Greek Literature*. Cambridge: Cambridge University Press, 2001.
- Roy, Ashley and Paul Robertson. "Applying Cosine Similarity to Paul's Letters: Mathematically Modeling Formal and Stylistic Similarities." In *New Approaches to Textual and Image Analysis in Early Jewish and Christian Studies*, edited by Garrick V. Allen, Sara Schulthess, Paul Dilley, and Peter Phillips, 88–117. Leiden: Brill, 2022.
- Savoy, Jacques. "Authorship of Pauline Epistles Revisited." *Journal of the Association for Information Science and Technology* 70 (2019), 1089–97.
- Schenk, Wolfgang. "Die Briefe und Timotheus I und Timotheus II und an Titus (Pastoralbriefe) in der neueren Forschung (1945–1985)." In *Aufstieg und Niedergang der römischen Welt: Geschichte und Kultur Roms im Spiegel der neueren Forschung*, edited by Hildegard Temporini and Wolfgang Haase, 3404–38. Berlin: de Gruyter, 1988.
- Schleiermacher, Friedrich. *Über den sogenannten Ersten Brief des Paulus an den Timotheus: Ein kritisches Senschreiben an J.C. Cass, Consistorialassessor und Feldprediger zu Stettin*. Berlin: De Gruyter, 2021. Originally published 1807.
- Spicq, Celes. *Les Épîtres Pastorales*. Paris: J. Gabalda et Cie, 1947.
- Stowers, Stanley K. "Friends and Enemies in the Politics of Heaven: Reading Theology in Philippians." In *Pauline Theology, Volume I: Thessalonians, Philippians, Galatians, Philemon*, edited by Joulette M. Bassler, 105–21. Minneapolis: Fortress, 2009.
- Stowers, Stanley K. *Letter Writing in Greco-Roman Antiquity*. LEC. Philadelphia: Westminster, 1986.
- Sullivan, Michael. *Statistics: Informed Decisions Using Data*. 4th ed. Pearson Education, Inc, 2013.
- Thraede, Klaus. *Grundzüge griechisch-römischer Brieftopik*. München: C.H. Beck, 1970.
- Torm, Frederick. "Über die Sprache in den Pastoralbriefen." *Zeitschrift für die Neutestamentliche Wissenschaft und die Kunde der Älteren Kirche* 18 (1918), 225–43.
- Van Nes, Jermo. *Pauline Language and the Pastoral Epistles: A Study of Linguistic Variation in the Corpus Paulinum*. LBS 16. Leiden: Brill, 2018.
- Williams, David S. "Josephus, Stylometry, and Jewish Studies." *SHOFAR* 11:4 (1993), 18–34.
- Wischmeyer, Oda (ed.). *Paul: Life, Setting, Work, Letters*. London: T&T Clark, 2012.
- Zipf, George Kingsley. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Hafner, 1972. Originally published 1949.

Appendix: Database of Stylistic Features

Romans

Declarations of fact: 1:2; 1:3; 1:16b–17; 1:18–32; 2:1–11; 2:12–16; 2:25–29; 3:21–26; 5:3b–5; 6:1–11; 6:23; 7:7–13; 7:14–25; 8:1–8; 8:9–11; 8:12–17; 8:18–25; 8:26–27; 8:28–30; 8:31–39; 9:4–5; 11:25–32; 12:4–8; 13:1–7; 13:8b–10; 13:11–12a; 14:4b; 14:6–9; 14:10b–12; 14:14–15; 14:17–23; 15:4; 15:8–12; 16:20

Appeals to authority: 1:9; 1:17; 2:24; 3:4; 3:10–18; 4:7–8; 4:17; 4:18; 4:22; 8:3b; 9:7; 9:9; 9:12; 9:13; 9:15; 9:17; 9:25–26; 9:27; 9:29; 10:5; 10:6–7; 10:8; 10:11; 10:13; 10:15; 10:16; 10:18; 10:19; 10:20; 10:21; 11:3–4; 11:26–27; 11:34–35; 12:19–20; 14:10b–11; 15:3–4; 15:9–12; 15:21

Logical arguments: 3:1–8; 3:9–10; 3:27–31; 4:1–8; 4:9–12; 4:13–25; 5:6–11; 5:12–14; 5:15–21; 7:1–6; 9:6–18; 9:19–29; 9:30–10:4; 10:5–17; 10:18–21; 11:1–6; 11:7–10; 11:11–16; 11:23–24

Inferences: 1:24–25; 2:1; 4:22; 5:1; 5:18; 6:4; 6:12; 7:12; 7:21; 7:25; 8:1; 8:12; 9:18; 10:17; 11:5–6; 13:2; 13:5; 13:10; 14:3; 14:12; 14:16; 14:19; 15:7; 15:17; 15:22

First-person self-fashioning: 1:1–5; 1:13–15; 1:16; 9:1–3; 11:1; 11:7–9; 11:13–14; 15:16–20; 15:22–29; 16:1

First-person emphases: 1:13; 2:16; 3:5; 6:19a; 7:1; 8:18; 8:38; 9:1–3; 10:18; 10:19; 11:1; 11:11; 11:13; 11:25; 12:1; 12:3; 14:14–15; 15:8; 15:15; 15:30; 16:1; 16:17

First-person figurations of groupness: 5:6–11; 6:5–9; 7:5–6; 8:12–17; 8:22–25; 8:37–39

Second-person claims about audience: 1:6; 1:7; 6:14; 6:16–23; 7:4; 8:9–11

Second-person emphases: 1:13; 7:1; 7:4; 8:12; 10:1; 11:25; 12:1; 12:19; 13:3

Third-person contrasts: 16:17–18

General exhortations or advice: 5:1–3; 6:11; 6:12–13; 6:19b; 11:18–20; 12:1; 12:2; 12:3; 12:9a; 12:9b; 12:10a; 12:10b; 12:11a; 12:11b; 12:11c; 12:12a; 12:12b; 12:12c; 12:13a; 12:13b; 12:14; 12:15; 12:16; 12:17; 12:18; 12:19; 12:20; 12:21; 13:1; 13:5; 13:6–7; 13:12; 13:13–14; 14:1; 14:3; 14:5; 14:13; 14:16; 14:19; 15:30–32; 16:17

Specific imperatives: 11:22; 14:15; 14:20; 14:22; 15:1–2; 15:7; 16:3–15; 16:16

Expressions of affection or friendship: 1:8–13; 7:1; 7:4; 8:12; 10:1; 11:25; 12:1; 12:19; 15:14; 15:22–29; 15:30–32; 16:17; 16:19

Expressions of confidence: 15:14

Second or third-person praise: 1:8; 16:19

Second-person blame: 3:8; 14:4; 14:5; 14:10

Satire:

Expressions of wish or desire: 1:7; 15:5–6; 15:13; 15:33; 16:20

1 Corinthians

Declarations of fact: 1:8; 1:18–31; 2:10–11; 2:14–16a; 3:11–15; 3:17; 3:19; 5:7b; 6:9–10; 8:1b–3; 10:1–11; 10:13; 10:23; 11:3–12; 11:23–26; 11:27; 11:29–32; 12:4–11; 12:12–13; 12:28–30; 13:5b–12; 13:13; 14:2–5; 14:32; 14:37–38; 15:3–8; 15:20–26; 15:36–49; 15:50–55; 15:56–57

Appeals to authority: 1:19; 1:31; 2:9; 2:16; 3:19b–20; 5:13; 6:16; 7:10; 9:8–10; 10:7; 10:26; 14:21; 15:32; 15:33; 15:54b–55

Logical arguments: 6:12–20; 8:4–13; 9:3–14; 10:15–22; 10:29b–30; 11:13–15; 11:33; 12:14–26; 14:6–12a; 14:14–17; 14:22–25; 15:12–19; 15:27–28; 15:29–33

Inferences: 1:7; 4:5; 4:17; 5:8; 7:38; 8:13; 9:26; 10:14; 10:31; 11:27; 12:3; 14:13; 14:23; 15:11; 15:58

First-person self-fashioning: 1:14–17; 1:23; 2:1–5; 2:6–7; 2:12–13; 2:16b; 3:5–9a; 3:10a; 4:1–4; 4:6; 4:9–13; 4:14; 4:19–21; 5:3–4a; 6:12; 7:25; 7:40b; 8:13; 9:1–14; 9:15–18; 9:19–23; 9:26–27; 10:31–11:1; 11:23; 12:31; 14:18–19; 15:3–8; 15:9–11a; 15:31–32; 16:3–4; 16:5–9; 16:10–11; 16:12; 16:17–18a

First-person emphases: 1:12; 4:14; 5:3; 5:9–11; 7:6; 7:8; 7:10; 7:12; 7:17; 7:25–26; 7:29; 7:32; 7:35; 9:3; 9:8; 9:15; 10:1; 11:3; 11:34b; 12:1; 12:3; 12:31; 14:5; 14:37; 15:1; 15:50; 15:51; 16:21

First-person figurations of groupness: 8:6; 9:25; 10:17; 11:16

Second-person claims about audience: 1:2; 1:5–7; 3:9b; 3:16–17; 3:21b–23; 6:11; 9:1b–2; 10:13; 10:15; 12:2; 12:27; 15:1b–2; 15:11b

Second-person emphases: 1:10; 1:11; 1:26; 2:1; 3:1; 4:6; 7:24; 7:29; 9:24; 10:1; 10:12; 10:15; 10:19–20a; 10:20b; 10:29; 11:33; 12:1; 14:6; 14:20; 14:26; 15:1; 15:31; 15:50; 15:58; 16:15

Third-person contrasts: 2:8; 15:34b

General exhortations or advice: 1:10; 3:10b; 4:1; 4:5; 4:17; 5:8; 5:11–12; 6:18; 6:20; 7:1–7; 7:8–9; 7:10–11; 7:12–16; 7:17–24; 7:25–31; 7:32–35; 7:36–38; 7:39–40a; 8:9; 9:24; 10:7; 10:8; 10:9; 10:10; 10:14; 10:24; 10:25; 10:27–28; 10:31–32; 11:1; 11:2–12; 11:28; 12:31; 14:1; 14:12; 14:13; 14:20; 14:26–33; 14:34–36; 14:37; 14:39–40; 16:10; 16:11a; 16:11b; 16:13a; 16:13b; 16:13c; 16:13d; 16:14; 16:16; 16:18b; 16:20b

Specific imperatives: 1:26; 3:18; 5:4b–5; 5:7; 5:13; 6:9; 10:15; 10:18; 11:13; 11:37; 15:33; 16:1–2; 16:20

Expressions of affection or friendship: 1:4–8; 1:10; 1:11; 3:1; 3:21; 4:6; 4:14–16; 10:1; 10:14; 11:33; 14:6; 14:20; 14:26; 14:39; 15:31; 15:50; 15:58; 16:7; 16:15

Expressions of confidence:

Second or third-person praise: 11:2

Second-person blame: 1:11–13; 3:1–4; 3:16; 4:7; 4:18–21; 5:1–2; 5:6; 6:1–6; 6:7–9; 6:15–19; 10:22; 11:17–22; 15:12; 15:34c

Satire: 4:8

Expressions of wish or desire: 1:3; 7:7; 16:22; 16:23; 16:24

2 Corinthians

Declarations of fact: 1:3; 3:13–18; 4:3–4; 9:6–15; 10:18; 13:3b–4

Appeals to authority: 1:23; 4:6; 4:13; 6:2; 6:16–18; 8:15; 9:9; 10:17; 11:31; 12:8–9; 13:1

Logical arguments: 3:7–11; 6:14–7:1; 7:11; 8:8–15

Inferences: 2:8; 3:12; 4:1; 4:12; 4:16; 5:11; 5:14b; 5:16; 7:12; 8:24; 10:18; 11:15; 12:7; 12:9b; 12:10

First-person self-fashioning: 1:1; 1:5–7; 1:8–11; 1:12–14; 1:15–2:4; 2:10–11; 2:12–13; 2:14–17; 3:4–6; 3:12; 4:1–6; 4:7–15; 4:16–5:10; 5:11–15; 5:16–21; 6:1–10; 7:2b; 7:5–13a; 7:13b; 8:20–21; 9:3–5; 10:3–6; 10:8–11; 10:12–18; 11:1–6; 11:12; 11:21b–33; 12:1–10; 13:4b; 13:10

First-person emphases: 1:8; 1:13; 1:23; 3:1a; 5:20; 6:1; 6:13; 7:3; 8:1; 8:8; 8:10; 9:1; 10:1a; 10:1b; 11:16; 11:17; 11:21a; 11:21b; 11:30; 11:31; 12:1; 12:19; 13:2; 13:6; 13:10

First-person figurations of groupness: 3:18a; 6:16–7:1

Second-person claims about audience: 3:11–13

Second-person emphases: 1:8; 3:1b–3; 6:11; 7:1; 8:1; 12:19; 13:2; 13:11

Third-person contrasts: 10:10–11; 10:12; 11:12–15; 11:20

General exhortations or advice: 5:20; 6:1; 6:14; 7:1; 8:10–11; 13:11c; 13:11d

Specific imperatives: 2:7–8; 6:13; 7:2; 7:11; 8:7; 8:24; 9:7; 10:1–2; 10:7a; 10:7b; 11:1b; 11:16; 12:14; 13:5; 13:11b; 13:11c; 13:12

Expressions of affection or friendship: 1:8; 1:11; 1:14; 2:4; 2:10; 3:1b–3; 6:11–13; 7:1; 7:3–4; 7:7–13a; 8:1; 8:16; 12:19; 13:9; 13:11

Expressions of confidence: 2:3; 7:16; 8:22

Second or third-person praise: 7:14–16; 8:1–6; 8:7; 8:16–24; 9:1–12

Second-person blame: 11:3–4; 11:7–11; 11:20; 12:11a; 12:14–18; 12:20–21; 13:1–3; 13:5

Satire: 11:5; 11:19; 11:21; 11:23; 12:11; 12:13b

Expressions of wish or desire: 1:2; 1:3; 11:11; 13:6–7; 13:9; 13:13

Galatians

Declarations of fact: 1:2; 1:4; 2:15–16; 3:2–5; 4:29; 5:6; 5:14; 5:17–18; 5:19–20; 5:22–23; 5:24; 6:3; 6:7–8; 6:15

Appeals to authority: 3:6; 3:8; 3:10; 3:11; 3:12; 3:13; 3:16; 4:23–26; 4:27; 4:30; 5:14

Logical arguments: 3:6–9; 3:10–14; 3:15–18; 3:19–20; 3:21–22; 4:1–6; 4:22–27

Inferences: 4:7; 4:31–5:1a; 5:1b; 6:10

First-person self-fashioning: 1:1–2; 1:10; 1:11–12; 1:13–14; 1:15–24; 2:1–2; 2:5; 2:6–9; 2:10; 2:11–14; 2:17–21; 3:2–4; 4:11; 4:12; 4:13–14; 5:11; 6:14; 6:17

First-person emphases: 1:9; 1:11; 1:20; 3:15; 3:17; 4:1; 4:15; 5:2–3; 5:16; 5:21

First-person figurations of groupness: 3:23–25; 4:3; 4:31; 5:5

Second-person claims about audience: 3:26–27; 3:29; 4:6–7; 4:8–9a; 4:28; 5:13

Second-person emphases: 1:11; 1:13; 3:1; 3:2–5; 3:15; 4:12; 4:13; 4:19; 4:21; 4:28; 4:31; 5:11; 5:13; 6:1; 6:18

Third-person contrasts: 1:7; 2:4; 2:12–14; 4:17–18; 5:10; 5:12; 6:12–13

General exhortations or advice: 4:12; 5:11; 5:13; 5:15; 5:16; 5:25; 5:26; 6:1b–2; 6:4–5; 6:6; 6:7a; 6:10; 6:17

Specific imperatives: 4:21; 4:30; 6:11

Expressions of affection or friendship: 1:11; 3:15; 4:12; 4:19; 4:31; 5:11; 5:13; 6:1; 6:18

Expressions of confidence: 5:10

Second or third-person praise:

Second-person blame: 1:6; 1:8–9; 3:2; 4:8–11; 4:14–16; 4:19–20; 4:21; 5:4; 5:7–8

Satire: 5:12

Expressions of wish or desire: 1:3; 3:2; 4:20; 6:16; 6:18

1 Thessalonians

Declarations of fact: 1:10b; 4:3–6; 4:7; 4:14–17; 5:2–3; 5:9–10; 5:24

Appeals to authority: 2:5; 2:10

Logical arguments:

Inferences: 3:1; 4:1; 4:8; 4:18; 5:6; 5:11

First-person self-fashioning: 2:1–2; 2:3–8; 2:9b; 2:10–12; 3:1–3a; 3:3b–4; 3:5

First-person emphases: 2:18; 4:6; 4:11; 4:14; 4:15

First-person figurations of groupness: 4:14; 5:5b; 5:9–10

Second-person claims about audience: 4:9; 5:4–5a

Second-person emphases: 1:4; 1:5; 2:1; 2:2; 2:5; 2:9a; 2:10; 2:11; 2:14; 2:17; 3:3; 3:4; 3:7; 4:1; 4:2; 4:10b; 4:13; 4:14; 5:1; 5:2; 5:12; 5:14

Third-person contrasts: 2:15–16; 5:2; 5:7

General exhortations or advice: 4:1; 4:10b–12; 5:6; 5:8; 5:12; 5:13a; 5:13b; 5:15; 5:16; 5:17; 5:18; 5:19; 5:20–21; 5:22; 5:25

Specific imperatives: 4:18; 5:11; 5:26; 5:27

Expressions of affection or friendship: 1:2–5; 2:1; 2:8; 2:9; 2:17–20; 3:6–10; 3:11–12; 4:1; 4:10; 5:1; 5:12; 5:14; 5:25

Expressions of confidence:

Second or third-person praise: 1:3; 1:6–10a; 2:13; 2:14; 3:6–8; 4:1; 4:9–10a

Second-person blame:

Satire:

Expressions of wish or desire: 1:1; 3:11; 5:23; 5:28

Philippians

Declarations of fact: 1:12–14; 1:15–18a; 1:28c–29; 2:6–8; 2:9–11; 2:13; 2:27; 3:21; 4:9

Appeals to authority: 1:8

Logical arguments:

Inferences: 2:1; 2:9; 2:12; 2:23; 2:28; 4:1

First-person self-fashioning: 1:1; 1:12–14; 1:15–18a; 1:18b–20; 1:21–26; 2:17–18; 2:19–23; 2:25–20; 3:4b–6; 3:7–9; 3:10–11; 3:12–14; 3:17; 3:18a; 4:1; 4:10; 4:11–13; 4:18

First-person emphases: 1:12; 3:1b; 3:18b; 4:2; 4:3; 4:4

First-person figurations of groupness: 3:3; 3:20–21

Second-person claims about audience: 2:15b

Second-person emphases: 1:2; 2:12; 3:1b; 3:13; 3:17; 4:1a; 4:1b; 4:15

Third-person contrasts: 1:15–17; 1:28b; 2:21; 3:2; 3:18–19

General exhortations or advice: 1:27–28a; 2:2; 2:3–4; 2:5; 2:12; 2:14; 2:18; 3:1; 3:2; 3:15a; 3:16; 3:17; 4:4; 4:5a; 4:6; 4:8; 4:9

Specific imperatives: 2:29; 4:2; 4:3; 4:21

Expressions of affection or friendship: 1:3–8; 1:12; 1:19; 1:26; 1:30; 2:12; 2:17–18; 2:19–23; 2:25–30; 3:1a; 3:13; 3:17; 4:1a; 4:1b; 4:10–14; 4:15–19

Expressions of confidence:

Second or third-person praise: 2:16; 2:20–22; 2:30
Second-person blame: 1:24–25; 2:30; 4:10
Satire: 4:10
Expressions of wish or desire: 1:2; 1:9–11; 4:20; 4:23

Philemon

Declarations of fact:
Appeals to authority:
Logical arguments: 15
Inferences:
First-person self-fashioning: 8–9; 10; 12–14
First-person emphases: 8; 9; 19a; 19b; 21
First-person figurations of groupness:
Second-person claims about audience:
Second-person emphases:
Third-person contrasts:
General exhortations or advice:
Specific imperatives: 17–18; 20; 22
Expressions of affection or friendship: 1; 4–7; 10; 12; 16; 20; 22
Expressions of confidence: 21
Second or third-person praise: 4–7
Second-person blame:
Satire:
Expressions of wish or desire: 3; 6; 25

1 Timothy

Declarations of fact: 1:5; 1:9–11a; 2:3–4; 2:15; 3:1; 3:13; 3:16; 4:4–5; 4:8–9; 5:4b–6; 5:8; 5:11b–13; 5:24–25; 6:6; 6:15–16
Appeals to authority: 2:5–6; 3:16; 4:1; 5:18
Logical arguments: 2:13–14; 3:5
Inferences: 2:1; 2:8
First-person self-fashioning: 1:1; 1:11b; 1:12–16; 1:20; 2:7; 3:14–15
First-person emphases: 1:3; 1:18–19; 2:1; 2:7; 3:14–15; 5:21; 6:13
First-person figurations of groupness: 6:7–9
Second-person claims about audience: 1:18; 4:6; 6:11; 6:12
Second-person emphases: 1:2; 1:18; 6:11; 6:20
Third-person contrasts: 1:6–7; 1:19b–20; 4:1–3; 5:15; 5:24a; 6:3–5; 6:9–10; 6:21
General exhortations or advice: 2:1–2; 2:11–12; 3:2–7; 3:8–10; 3:11; 3:12; 4:7a; 4:7b; 4:11; 4:12; 4:13; 4:14; 4:15; 4:16; 5:1–2; 5:3–4a; 5:7; 5:9–16; 5:17–18; 5:19; 5:20; 5:21; 5:22a; 5:22b; 6:1–2a; 6:2b; 6:11; 6:12a; 6:12b; 6:13–14; 6:17–18; 6:20a; 6:20b
Specific imperatives: 1:3–4
Expressions of affection or friendship: 1:2; 1:18; 3:14
Expressions of confidence:
Second or third-person praise:
Second-person blame:
Satire:
Expressions of wish or desire: 1:2; 1:17; 2:8; 2:9; 6:16b; 6:20b

2 Timothy

Declarations of fact: 1:7; 1:9–10; 2:11–13; 2:20–21; 3:1–9; 3:12–13; 3:16–17; 4:17–18
Appeals to authority: 2:11–13; 2:19

Logical arguments: 2:4–6

Inferences: 1:8

First-person self-fashioning: 1:1; 1:3; 1:11–12; 2:8b–10; 3:10–11; 4:6–8

First-person emphases: 1:5; 1:6; 4:1; 4:6

First-person figurations of groupness:

Second-person claims about audience: 3:10; 3:14–15

Second-person emphases: 1:1; 1:15; 1:18; 2:1; 3:14

Third-person contrasts: 1:15; 2:17b–18; 2:25b–26; 3:1–9; 3:13; 4:3–4; 4:14–15

General exhortations or advice: 1:6; 1:8; 1:13–14; 2:1; 2:2; 2:3; 2:8; 2:14; 2:15; 2:16–17a; 2:22; 2:23; 2:24–25a; 3:1; 3:5; 3:14–15; 4:2a; 4:2b; 4:2c; 4:5a; 4:5b; 4:5c; 4:5d; 4:15

Specific imperatives: 2:7; 4:9; 4:11; 4:13; 4:19; 4:21

Expressions of affection or friendship: 1:1; 1:3–5; 2:1

Expressions of confidence:

Second or third-person praise: 1:5

Second-person blame:

Satire:

Expressions of wish or desire: 1:2; 1:16–18; 4:16; 4:18b; 4:22

Titus

Declarations of fact: 1:5; 2:11–14; 3:8a; 3:8b

Appeals to authority:

Logical arguments:

Inferences:

First-person self-fashioning: 1:1–3

First-person emphases: 1:5

First-person figurations of groupness: 3:3–7

Second-person claims about audience:

Second-person emphases: 1:4; 2:1; 3:11

Third-person contrasts: 1:10–16

General exhortations or advice: 1:5b–9; 1:11; 1:13; 2:1; 2:2; 2:3–5; 2:6; 2:7–8; 2:9–10; 2:15a; 2:15b; 2:15c; 3:1–2; 3:9–10; 3:14

Specific imperatives: 3:12; 3:13; 3:15

Expressions of affection or friendship: 1:4

Expressions of confidence:

Second or third-person praise:

Second-person blame:

Satire:

Expressions of wish or desire: 1:4; 3:8; 3:15