Digital Scholarship in the Humanities, 2023, **38**, 647–657
https://doi.org/10.1093/llc/fqac078
Advance access publication 19 December 2022
Short Paper

OXFORD

Likelihood calculations for reconstructed lacunae and Papyrus 46's text of Ephesians 6:19

Joey McCollum *[1]

[1]Institute for Religion and Critical Inquiry, Australian Catholic University, Melbourne, VIC, Australia

*Correspondence: Joey McCollum, Institute for Religion and Critical Inquiry, Australian Catholic University, Melbourne, VIC, Australia.
E-mail: james.mccollum@myacu.edu.au

Abstract

This article presents an approach to quantify the likelihood of a given reconstruction of lacunose text in a manuscript using statistics on line lengths (in letters), information about the line-breaking conventions and scribal habits of the scribe who copied the manuscript, and the well-known computational technique of dynamic programming. The approach and its value are illustrated with an application to a textual contest between the readings τὸ μυστήριον τοῦ εὐαγγελίου and τὸ μυστήριον in Ephesians 6:19, where the early papyrus witness P. Chester Beatty II/P. Mich. Inv. 6238 (Gregory–Aland 𝔓[46]) is lacunose. The study shows that under reasonable assumptions, 𝔓[46] is over fifty times more likely to have read τὸ μυστήριον.

1 Motivation

The Greek text of Ephesians 6:19 features a textual variant that has defied an obvious solution on the basis of the intrinsic qualities of the variant readings themselves. In context, the author of the letter is asking his readers to pray for him so that he might be able to preach boldly the message for which he serves as an imprisoned ambassador. The variant concerns the identity of this message: it could be τὸ μυστήριον τοῦ εὐαγγελίου, 'the mystery of the gospel', or simply τὸ μυστήριον, 'the mystery'.

While the longer reading is less obscure than the shorter one, this may be because it arose as a clarification of the shorter reading. Conversely, while the shorter reading's meaning is discernable from context and could be authorial, it could also be a later harmonization to the (at least superficially) unqualified use of the same phrase in Ephesians 3:3, 9. These and other arguments can be found in the literature and will not be rehearsed here (Weiss, 1896, p. 89; Schnackenburg, 1991, p. 283, n. 53; Best, 1998, p. 608); it suffices to note that commentators who do not resolve this crux on the basis of the readings have turned to the preponderance of manuscript evidence to do so (Eadie, 1861, p. 489; Ellicott, 1884, p. 155; Bruce, 2012, p. 108; Hoehner, 2002, p. 860, n. 3).

But even the external evidence leaves room for doubt. In favor of the shorter reading, the fourth-century codex Vaticanus, hereafter denoted by the Gregory–Aland (GA) number 03 or the traditional siglum B, joins an otherwise 'Western' bloc of witnesses consisting of the later Greek–Latin diglots GA 010 (=F) and 012 (=G), a few manuscripts of the Old Latin tradition, and quotations from some Latin-speaking Church fathers. Metzger's commentary is worth quoting in this connection:

> Although it may appear noteworthy that B joins it[g, mon] *al* in supporting the shorter reading, in the Pauline corpus codex Vaticanus not infrequently displays a strand of Western contamination, and therefore the weight of its testimony, when united with Western witnesses, should not be overevaluated. (Metzger, 1994, p. 542)

This echoes the earlier judgment of Hort that in the Pauline Epistles there is an unquestionable intermingling of readings [in Codex Vaticanus] derived from a Western text nearly related to that of G_3 [=012] (Westcott and Hort, 1881–1882, 2:150). Thus, without some other early non-Western witness to support it, the shorter reading is left to be dismissed as an essentially isolated Western reading.

Under normal circumstances, the papyrus P. Chester Beatty II/P. Mich. Inv. 6238 (hereafter 𝔓[46]), commonly dated to the third century, might serve as such a witness. The papyrus has been the subject of close study for decades and is considered as one of our most important textual witnesses to the corpus of Paul's letters.[1] It is known to share many early readings with GA 03 (Zuntz, 1953, pp. 39–41, 61–68), so an agreement between $\mathfrak{P}^{46}$ and 03 here would add compelling diversity and antiquity to the shorter reading.

This makes the lacuna of the last three lines of fol. 80^{v} of $\mathfrak{P}^{46}$ (a common problem throughout the papyrus as a result of its having frayed at the bottom) all the more painful a loss. It becomes clear from a comparison of Metzger's assessment and the judgments of scholars who count $\mathfrak{P}^{46}$ among the support for the shorter reading that the 'preponderance' of the manuscript evidence hinges on whether $\mathfrak{P}^{46}$ read τὸ μυστήριον τοῦ εὐαγγελίου or τὸ μυστήριον. Given the weight he assigns to the papyrus, it is likely that Zuntz favors the shorter reading precisely because he thinks $\mathfrak{P}^{46}$ supports it: 'The relevant line of the papyrus is indeed missing. H. C. Hoskier observed that considerations of space suggest the shorter reading (*J.T.S.* xxxviii, 1937, 158; *Appendix* …, p. 14); he adds "Sir Frederic Kenyon agrees that the omission is almost certain"' (Zuntz, 1953, p. 95, n. 1). Comfort follows the same reasoning when he judges τὸ μυστήριον 'the original reading according to the two earliest manuscripts ($\mathfrak{P}^{46\ vid}$ B)'. He claims that 'The reconstruction of the lines in $\mathfrak{P}^{46}$ allow for "mystery" to fit the space as opposed to "mystery of the gospel"' (Comfort, 2015, p. 346). Unfortunately, he does not offer such a reconstruction, and in the latest edition of their transcription of $\mathfrak{P}^{46}$, he and Barrett make no attempt at one (Comfort and Barrett, 2019, p. 287).

The two latest transcriptions of its text reconstruct the missing lines, but they conflict with each other (Ebojo, 2014; Peterson, 2020). Ebojo, on the one hand, reconstructs the last three lines of fol. 80^{v} to include the longer reading (Ebojo, 2014, p. 798). The reconstruction (with supplied lacunose text in brackets and dots under unclear letters) is included below, with line lengths in letters on the right and the preceding line included for context:

[σει και δεη]σ̣ε̣ι̣ π̣[ερι παντων τω]ν̣α̣[γιων και]	(33)
[υπερ εμου ινα μοι δοθη λογος εν ανοιξει του στο]	(38)
[ματος μου εν παρρησια γνωρισαι το μυστηριον]	(37)
[του ευαγγελιου υπερ ου πρεσβευω εν αλυσει]	(35)

Peterson's reconstruction, on the other hand, assumes the shorter reading and produces lines more consistent with the average line length on this page (Peterson, 2020, p. 581):

[σει και δεησει περι παντων τω]ν α̣γ[ιων και]	(33)
[υπερ εμου ινα μοι δοθη λογος εν ανοιξει]	(32)
[του στοματος μου εν παρρησια γνωρισαι]	(32)
[το μυστηριον υπερ ου πρεσβευω εν αλυσει]	(33)

Both readings are therefore possible. But which is more probable? To find out, we need a way to evaluate numerically the probability that a lost line would be of a certain length.

2 Method

This section will detail how we can evaluate the desired probabilities. The subsection on the basic model used to do this (Section 2.1) assumes some familiarity with common concepts from statistics such as populations and their parameters (specifically, the mean, variance, and standard deviation), probability distributions (and the Gaussian or normal distribution in particular), and statistical independence.[2] Some concepts familiar to computer scientists and programmers appear in Section 2.2, but I will illustrate and explain these in that section.

2.1 Modeling probabilities of line lengths

A natural way to model the likelihoods of line lengths in $\mathfrak{P}^{46}$ is with a probability distribution. Such a distribution is governed by measurable statistics from the papyrus—specifically, the mean length of the extant lines in the papyrus and the variance of line lengths around this mean. This model is clearly suitable at the level of an individual page, because the scribe's tendency to set an intended line length for each page on the first line, noted by Ebojo (2014, pp. 114–18), suggests a well-defined mean around which other line lengths will naturally cluster. But fol. 80^{v} contains only twenty-eight lines whose lengths we can discern with certainty, and such a small population is too sensitive to variations.

To average out any outlying line lengths and sources of variance like different widths of letters (117–118), defects in the papyrus offsetting subsequent lines (284–285), and imperfect justification at the ends of lines (286), it is helpful to extend the population to the lines of other folios, excepting title lines and closing lines (since these are typically not complete) and insufficiently extant lines.[3] At the same time, we must also be wary of trends involving changes in line length over the course of the papyrus.[4] To this end, we will limit our population of lines to the folios of $\mathfrak{P}^{46}$ that contain Ephesians. Despite small variations from page to page, the distribution of letters in these folios tends to be tight and consistent in practice, as shown in Fig. 1.

The mean line length (given by the line in each box) ranges from thirty to thirty-five letters on every page of Ephesians except 75^{v} (where it is about twenty-eight letters) and 77^{v} (just above twenty-nine). These slightly

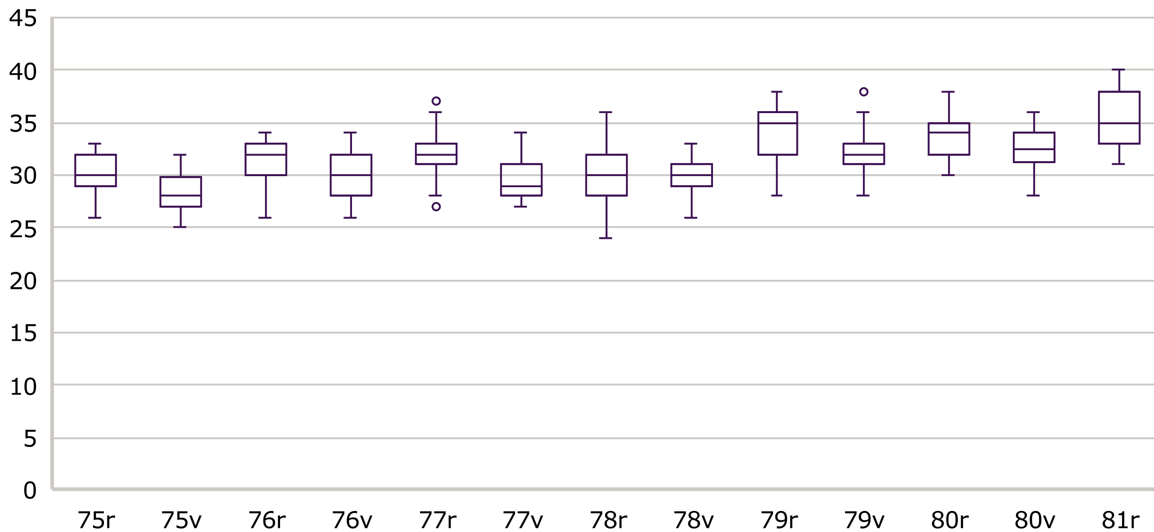

Figure 1. Box-and-whisker plot of line lengths in the pages of 𝔓[46] containing Ephesians

lower means are consistent with the trend, observed by Sanders and confirmed by Ebojo, that the lines on the recto pages of 𝔓[46] are often longer than those on its verso pages (Sanders, 1935, p. 6; Ebojo, 2014, p. 115). The line lengths in the inner quartiles for each page (represented by the boxes) all fall within a similar span of five letters, giving us an idea of the usual variation in line length due to the factors described above.

Taking these data as our extended population, we get $N = 330$ lines with a mean length of $\mu \approx 31.5576$ letters and a standard deviation of $\sigma \approx 2.9011$ letters. A natural probability distribution for our data is the normal distribution with this mean and standard deviation (hereafter denoted $N(x, \mu, \sigma)$, where x represents a line length), which corresponds to a bell curve like the one fit over the histogram of our line lengths in Fig. 2.

A more objective criterion for the quality of the fit is the coefficient of determination, denoted R^2, which measures the proportion of variance in the data explained by the model. This coefficient takes a value between 0 and 1, with a value close to 1 indicating that the model accounts for all the variance. For our population, this is $R^2 \approx 0.9288$, which means that roughly 7% of the variation in our data is unexplained. As shown in Fig. 2, most of this unexplained variance consists of an excess of lines twenty-eight letters long and a deficit of lines thirty-four letters long. The factors mentioned above might account for this, as might other considerations of the scribe's aesthetic preferences on when and how to break words at the end of a line. Our R^2 value also reassures us that our sample size of 330 lines is sufficient to accommodate the resulting variation. If our sample size had been too small, then the contribution of outlying line lengths would be influential enough to bring the R^2 value much closer to 0.

The normal distribution is a continuous distribution, meaning that it can be evaluated for any real number. If the line length were a continuous variable (e.g. length in centimeters), then we would simply compute the probability of a line with length x as $\Pr[X = x] = N(x, \mu, \sigma)$. (Here, X represents the variable for line length, while x represents a specific line length whose probability we want to know.) But since line length in letters is a discrete variable (i.e. since we assume that line lengths do not involve fractions of letters), it is better to use the continuity correction

$$\Pr[X = x] = \int_{x-0.5}^{x+0.5} N(X, \mu, \sigma)\mathrm{d}X \tag{1}$$

to evaluate the normal distribution as if it were a discrete distribution. The notation in Equation (1) means that we measure the area under the bell curve for one unit on the X-axis centered at the target value x—in other words, we treat the area in the unit as a discrete 'slot' of probability in the distribution. To return to the example of 𝔓[46], if we wanted to know the probability of the final thirty-five letter line in Ebojo's reconstructions of the lacuna in Ephesians 6:19, we would evaluate the area under the bell curve from $X = 34.5$ to $X = 35.5$, and this would give us a probability of $\Pr[X = 35] \approx 0.0682$. This is illustrated in Fig. 3.

Because we assume in our model of line length distributions that line lengths are independent variables drawn from the same normal distribution,[5] the probability of multiple lines with given lengths is simply the product of their individual probabilities. So, the probability of Ebojo's reconstruction of the last three lines on fol. 80$^{\mathrm{v}}$ with the longer reading can be calculated as

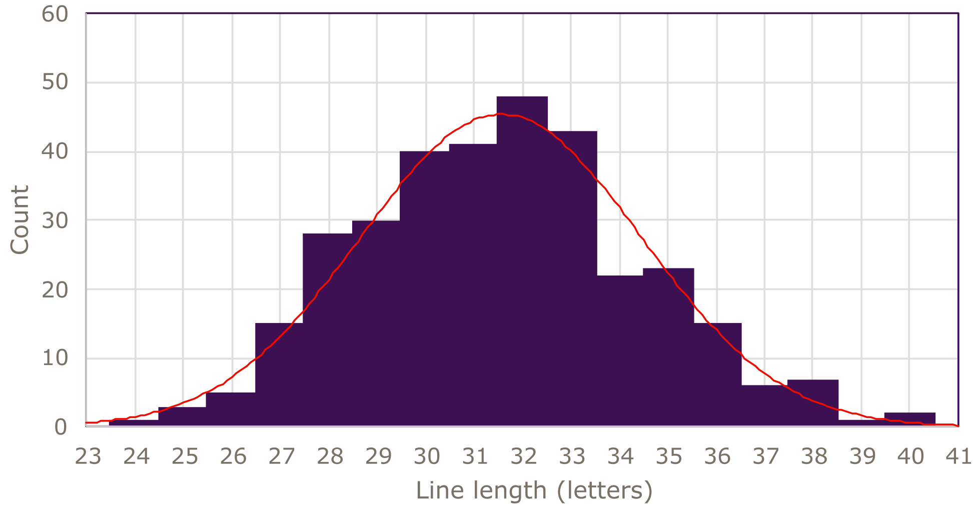

Figure 2. Histogram and corresponding scaled normal distribution (scaled by the population size N) of line lengths on all pages containing Ephesians in $\mathfrak{P}^{46}$

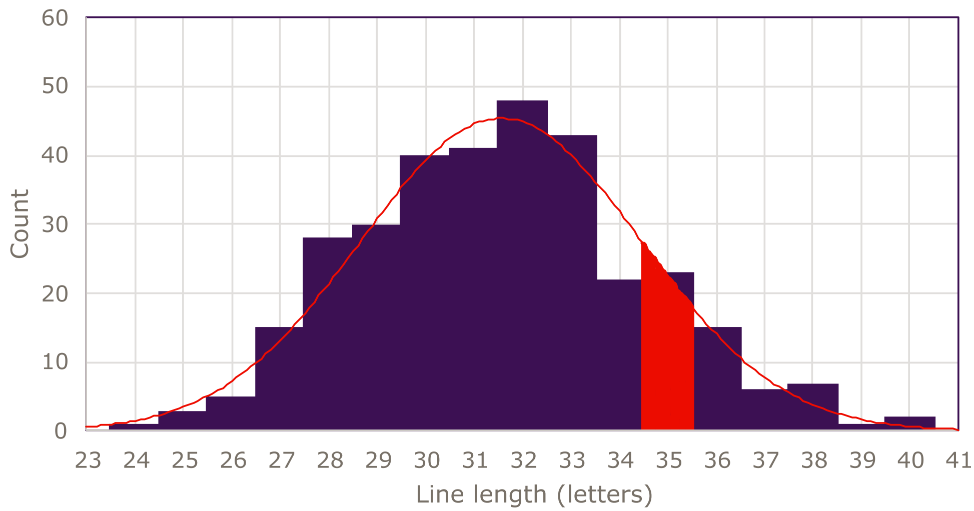

Figure 3. Approximation of a discrete probability with the normal distribution using the continuity correction

$$\begin{aligned}\Pr[X=38]\Pr[X=37]\Pr[X=35] &= \int_{37.5}^{38.5} N(X,\mu,\sigma)\mathrm{d}X \\ &\times \int_{36.5}^{37.5} N(X,\mu,\sigma)\mathrm{d}X \\ &\times \int_{34.5}^{35.5} N(X,\mu,\sigma)\mathrm{d}X \\ &\approx 1.9446\times 10^{-5},\end{aligned}$$

while the probability of Peterson's reconstruction with the shorter reading is

$$\begin{aligned}\Pr[X=32]\Pr[X=32]\Pr[X=33] &= \int_{31.5}^{32.5} N(X,\mu,\sigma)\mathrm{d}X \\ &\times \int_{31.5}^{32.5} N(X,\mu,\sigma)\mathrm{d}X \\ &\times \int_{32.5}^{33.5} N(X,\mu,\sigma)\mathrm{d}X \\ &= 0.0022.\end{aligned}$$

In other words, Peterson's reconstruction is over 113 times more likely according to this simple model.

2.2 Accounting for all segmentations

Of course, other reconstructions of the missing lines are possible. When we want to calculate the probability that $\mathfrak{P}^{46}$ had one reading or the other in a three-line lacuna, what we really want to calculate is the probability that it had this reading in *any* plausible reconstruction of the lacunose text that can be divided over three lines. A naïve way to do this would be to take the total number of letters in the reconstructed text and add up the probabilities of all possible partitions of these letters into three lines. For Ebojo's reconstruction of the three lost lines, which has a total of 113 letters, this calculation would look like

$$\sum_{\substack{i,j,k \geq 1 \\ i+j+k=113}} \Pr[X=i]\Pr[X=j]\Pr[X=k],$$

where the notation on the left indicates the sum over all line lengths i, j, and k that total 113 and the probabilities being added in this sum are calculated according to Equation (1).

But not every partition of a given number of letters across a given number of lines is valid according to the conventions of Greek line-breaking that $\mathfrak{P}^{46}$ follows, so we would like to restrict our space of possibilities to the valid ones. How can we carry out a computation of the probability subject to such qualifications? We can start by adding potential breakpoints to our candidate reconstructions according to the manuscript's line-breaking conventions. To give a short, simple example, the phrase ὑπὲρ ἐμοῦ would be encoded with the standard breakpoints as υ-περ ε-μου. The syllables of a textual sequence separated by hyphens or spaces in this way are called the 'tokens' of that text.

At its core, our task is to calculate the total probability of every way we can divide a sequence of tokens into a given number of contiguous segments. At first glance, this may not seem to make the problem easier, but in fact, the problem of sequence segmentation is known to admit an efficient solution using a tactic known as dynamic programming (Bellman, 1961; Cormen *et al.*, 2009, pp. 359–413). This approach is illustrated in Fig. 4.

It is helpful to think of the text and its possible divisions into lines as a graph, where the nodes represent the tokens and an edge from node i to node j indicates that a line contains everything after token i and up to and including token j. To ensure that the first token is not excluded, we add a placeholder token (denoted by asterisks in Fig. 4) with index 0 for the end of the previous line. No edges are allowed from a node to itself (because we assume that there are no empty lines), and all edges must point forward (because we cannot go backwards to break a line). So, in Fig. 4, the red edges represent the division υ/περ εμου, the blue edges represent υπερ/εμου, and the green edges represent υπερ ε/μου. Their weights (i.e. the labels on their edges in Fig. 4) correspond to the probability of a line containing all the letters that occur after the edge's starting token and up to and including its destination token, calculated according to the model in Section 2.1. So, the weight $w_{0,2}$ corresponds to the probability of a four-letter line (i.e. one consisting of υπερ) based on a normal distribution of line lengths.

The crucial component of this approach is the dynamic programming (hereafter denoted by DP) table. It is helpful to visualize this table as having rows numbered from 0 to n (inclusive) and columns numbered from 0 to L (inclusive), where n denotes the number of tokens and L denotes the desired number of lines. Thus, the table has $(n+1)\times(L+1)$ entries in total. The first entry of the table, corresponding to having a path of length 0 at the asterisk placeholder token, is initialized at 1.0, or a probability of 100%. The probability that we want to calculate for all segmentations of the text up to token n into L lines will be in the last entry of the table, which we will denote DP[n][L]. The algorithm for filling up the table is outlined in Algorithm

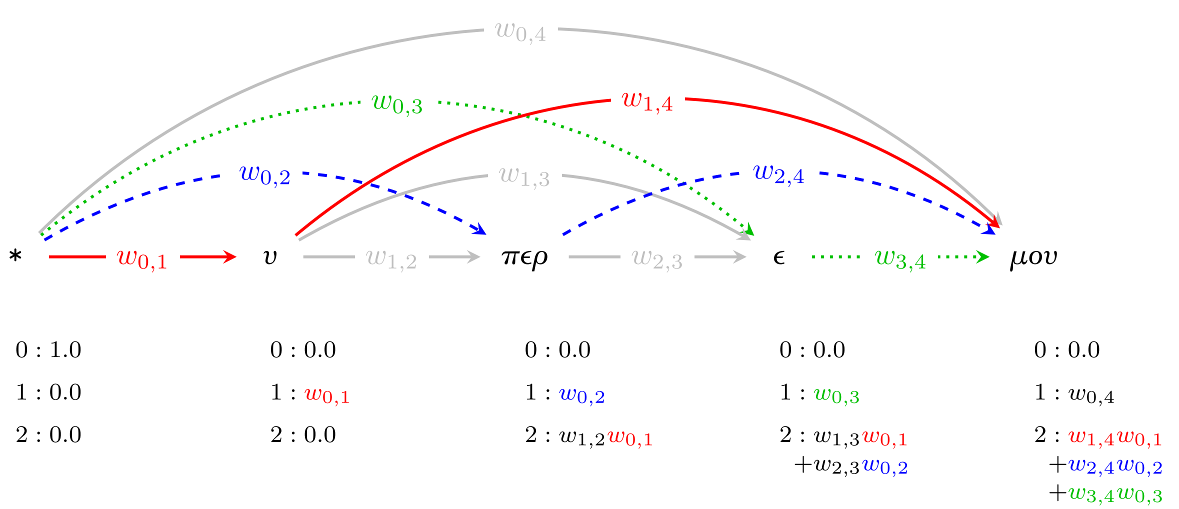

Figure 4. Illustration of dynamic programming applied to the segmentation problem of breaking the sequence υ-περ ε-μου into two lines

Algorithm 1 Pseudocode for filling a dynamic programming table with total probabilities of segmentations

```
Require: DP, n, L                                ▷ DP is an (n+1) × (L+1) table with all entries set to 0
DP[0][0] ← 1.0                                   ▷ Start with a probability of 100%
for i ← 0, n − 1 do                              ▷ Token at end of previous line
    for j ← i + 1, n do                          ▷ Token at end of current line
        for k ← 0, L − 1 do                      ▷ Index of previous line
            DP[j][k + 1] ← DP[j][k + 1] + (w_{i,j} × DP[i][k])   ▷ Update:
  Add probability of all segmentations in which letters after token i and up to token j occur on next line
        end for
    end for
end for
return DP[n][L]
```

1.If we proceed in this way for each node in order, then by the time we get to any node that follows it, that node's entries in the DP table will have already been fully populated. Thus, by the end of this process, the table entry $DP[n][L]$ will contain the desired probability of all divisions of the n-token text into L lines.

Because it recycles previous computations instead of making them fresh each time, DP is substantially faster than the brute-force approach of checking each feasible division of the text and calculating its probability. For n tokens and L desired lines, the DP algorithm takes on the order of n^2L operations (e.g. additions, multiplications) to terminate, while the brute-force approach takes on the order of n^L operations. To give a sense of the difference, for a text with $n = 50$ tokens to be segmented across $L = 4$ lines, DP is over six hundred times faster.

3 Validation

We are now prepared to apply the approach just outlined to the lacuna in Ephesians 6:19–20. But before we address the more contested matter of Ephesians 6:19, it will be useful to demonstrate the validity of our approach using reconstructions of lacunose passages in $\mathfrak{P}^{46}$ on which Ebojo and Peterson agree. In particular, we must consider passages where (1) a variant reading not attested by $\mathfrak{P}^{46}$ occurs and (2) that variant reading effects a difference in the length of the reconstructed text. Such passages are rare, but there are two that we can consider before we proceed to Ephesians 6:19.

3.1 Ephesians 2:20

Two lines containing text from Ephesians 2:20–22 are lost at the bottom of fol. 76ᵛ of $\mathfrak{P}^{46}$. Both Ebojo and Peterson reconstruct the missing text using the following sequence:

$\mathscr{T}_1$ = των α-πο-στο-λων και προ-φη-των ον-τος α-κρο-γω-νι-αι-ου αυ-του χρυ ιηυ εν ω πα-σα οι-κο-δο-μη.

But here, GA 010 and 012, the first hands of 06 and 0319, and the minuscule 1751 add λίθου after ἀκρογωνιαίου and transpose the order of the *nomina sacra* $\overline{\chi\rho\upsilon}$ $\overline{\iota\eta\upsilon}$ to $\overline{\iota\eta\upsilon}$ $\overline{\chi\rho\upsilon}$, yielding the following sequence:

$\mathscr{T}_2$ = των α-πο-στο-λων και προ-φη-των ον-τος α-κρο-γω-νι-αι-ου λι-θου αυ-του ιηυ χρυ εν ω πα-σα οι-κο-δο-μη.

Applying our approach to both sequences with the same mean and standard deviation for line lengths, we get the following probabilities for the sequences fitting into four lines: $\Pr[\mathscr{T}_1, L = 2] \approx 0.0350$ for Ebojo and Peterson's reconstruction, and $\Pr[\mathscr{T}_2, L = 2] \approx 0.0077$ for the reconstruction with the variant reading. The transposition of two *nomina sacra* of the same length obviously has no effect on the probability of the second reading, but the additional five letters of λίθου lower its probability considerably. Ebojo and Peterson's judgment, at least concerning the variant in question, is confirmed: their reconstruction is over four-and-a-half times more likely than the alternative.

3.2 Ephesians 4:15

Four lines containing text from Ephesians 4:15–16 are almost entirely lost at the bottom of fol. 78ʳ of $\mathfrak{P}^{46}$. Both Ebojo and Peterson reconstruct the missing text using the following sequence (where μενοι continues the περιφερόμενοι started on the previous line):

$\mathscr{T}_1$ = με-νοι παν-τι α-νε-μω της δι-δασ-κα-λι-ας εν τη κυ-βει-α των αν-θρω-πων εν παν-ουρ-γι-α προς την με-θο-δει-αν της πλα-νης α-λη-θευ-ον-τες δε εν α-γα-πη αυ-ξη-σω-μεν εις αυ-τον.

But the substitution of ἀλήθειαν δὲ ποιοῦντες for ἀληθεύοντες δέ, which finds early attestation in GA 010 and 012, the Vulgate, early Latin manuscripts and patristic citations, and is possibly supported outside the Latin-speaking world by the Bohairic Coptic version (Dubois, 1974), is also possible, which yields the following alternative sequence:

$\mathscr{T}_2$ = με-νοι παν-τι α-νε-μω της δι-δασ-κα-λι-ας εν τη κυ-βει-α των αν-θρω-πων εν παν-ουρ-γι-α προς την με-θο-δει-αν της πλα-νης α-λη-θει-αν δε ποι-ουν-τες εν α-γα-πη αυ-ξη-σω-μεν εις αυ-τον.

Applying our approach to both sequences with the same mean and standard deviation for line lengths, we get the following probabilities for the sequences fitting into four lines: $\Pr[\mathscr{T}_1, L = 4] \approx 0.0049$ for the reconstruction with ἀληθεύοντες δέ, and $\Pr[\mathscr{T}_2, L = 4] \approx 0.0041$ for that with ἀλήθειαν δὲ ποιοῦντες. The six additional letters of the second reading tilt the scales only slightly in favor of the first reading, but our approach nevertheless confirms Ebojo and Peterson's judgment regarding this variant.

3.3 Ephesians 6:19

Finally, we return to the variant at hand. In the three lacunose lines on fol. 80^{v} of $\mathfrak{P}^{46}$, Ebojo's reconstruction of Ephesians 6:19–20 with the longer reading is encoded with word breakpoints as

$\mathscr{T}_1$ = υ-περ ε-μου ι-να μοι δο-θη λο-γος εν α-νοι-ξει του στο-μα-τος μου εν παρ-ρη-σι-α γνω-ρι-σαι το μυ-στη-ρι-ον του ευ-αγ-γε-λι-ου υ-περ ου πρε-σβευ-ω εν α-λυ-σει,

while Peterson's reconstruction with the shorter reading is encoded as

$\mathscr{T}_2$ = υ-περ ε-μου ι-να μοι δο-θη λο-γος εν α-νοι-ξει του στο-μα-τος μου εν παρ-ρη-σι-α γνω-ρι-σαι το μυ-στη-ρι-ον υ-περ ου πρε-σβευ-ω εν α-λυ-σει.

Using the DP algorithm described in the previous section, we calculate the probability of the sequence $\mathscr{T}_1$ occurring over three lines is $\Pr[\mathscr{T}_1, L = 3] \approx 0.0001$. The corresponding probability for the sequence $\mathscr{T}_2$ with the shorter reading is $\Pr[\mathscr{T}_2, L = 3] \approx 0.0103$. So, under this assumption, it is over 100 times more likely that $\mathfrak{P}^{46}$ read τὸ μυστήριον than that it read τὸ μυστήριον τοῦ εὐαγγελίου. Furthermore, if we assume that $\mathfrak{P}^{46}$ must have had one of these two reconstructed texts, then the probabilities given this assumption can be expressed as

$$\begin{aligned}\Pr[\text{reading 1} \mid \mathscr{T}_1 \text{ or } \mathscr{T}_2] &= \frac{\Pr[\mathscr{T}_1, L=3]}{\Pr[\mathscr{T}_1, L=3] + \Pr[\mathscr{T}_2, L=3]} \\ &\approx \frac{0.0001}{0.0001 + 0.0103} \\ &\approx 0.0096\end{aligned}$$

and

$$\begin{aligned}\Pr[\text{reading 2} \mid \mathscr{T}_1 \text{ or } \mathscr{T}_2] &= \frac{\Pr[\mathscr{T}_2, L=3]}{\Pr[\mathscr{T}_1, L=3] + \Pr[\mathscr{T}_2, L=3]} \\ &\approx \frac{0.0103}{0.0001 + 0.0103} \\ &\approx 0.9904\end{aligned}$$

In comparison with the examples where Ebojo and Peterson agree on the reconstruction, our results for Ephesians 6:19 point strongly in one direction. The wider gap in probabilities is due to the more substantial difference in the length of the reconstructions: τοῦ εὐαγγελίου is a full thirteen letters, which exceeds the combined difference in length for our other two examples discussed above and has a more obvious effect over just three lines of text.

4 Limitations

Obviously, while the proposed approach can discern how many letters most likely occurred in a lacuna, it cannot make such determinations about what those letters were. In practical terms, this means that the method is suitable for weighing pluses and minuses in the text (especially when the difference in length between the readings is large, as it is in the example discussed here), but not suitable for weighing substitutions or transpositions.

As an example, the first hand of the fifteenth-century minuscule GA 636 has the reading τὸ εὐαγγέλιον. If we denote by $\mathscr{T}_3$ the reconstructed three-line lacuna containing this reading, then by the DP approach, then the probability of this text occurring over the three lines in question is $\Pr[\mathscr{T}_3, L = 3] \approx 0.0098$, which is very close to the probability of the same text with τὸ μυστήριον in the lacuna, $\Pr[\mathscr{T}_3, L = 3] \approx 0.0103$. As suggested above, this is because the two shorter readings are nearly the same length. If we assume that $\mathscr{T}_1, \mathscr{T}_2$, and $\mathscr{T}_3$ are the only possible sequences, and we use the probabilities that we calculated for those sequences in Section 3, then the probability that $\mathfrak{P}^{46}$ read τὸ μυστήριον τοῦ εὐαγγελίου becomes

$$
\begin{aligned}
&\Pr[\text{reading } 1 \mid \mathscr{T}_1 \text{ or } \mathscr{T}_2 \text{ or } \mathscr{T}_3] \\
&= \frac{\Pr[\mathscr{T}_1, L=3]}{\Pr[\mathscr{T}_1, L=3] + \Pr[\mathscr{T}_2, L=3] + \Pr[\mathscr{T}_3, L=3]} \\
&\approx \frac{0.0001}{0.0001 + 0.0103 + 0.0098} \\
&\approx 0.0063,
\end{aligned}
$$

the probability that it read τὸ μυστήριον becomes

$$
\begin{aligned}
&\Pr[\text{reading } 2 \mid \mathscr{T}_1 \text{ or } \mathscr{T}_2 \text{ or } \mathscr{T}_3] \\
&= \frac{\Pr[\mathscr{T}_2, L=3]}{\Pr[\mathscr{T}_1, L=3] + \Pr[\mathscr{T}_2, L=3] + \Pr[\mathscr{T}_3, L=3]} \\
&\approx \frac{0.0103}{0.0001 + 0.0103 + 0.0098} \\
&\approx 0.5104,
\end{aligned}
$$

and the probability that it read τὸ εὐαγγέλιον becomes

$$
\begin{aligned}
&\Pr[\text{reading } 3 \mid \mathscr{T}_1 \text{ or } \mathscr{T}_2 \text{ or } \mathscr{T}_3] \\
&= \frac{\Pr[\mathscr{T}_3, L=3]}{\Pr[\mathscr{T}_1, L=3] + \Pr[\mathscr{T}_2, L=3] + \Pr[\mathscr{T}_3, L=3]} \\
&\approx \frac{0.0098}{0.0001 + 0.0103 + 0.0098} \\
&\approx 0.4832.
\end{aligned}
$$

So, the inclusion of τὸ εὐαγγέλιον effectively 'splits the vote' that would have gone to τὸ μυστήριον.

Besides demonstrating a limitation of the proposed approach, this example emphasizes the necessity for critical judgment in our choice of inputs. While τὸ εὐαγγέλιον is perfectly fitting as an authorial reading on its own merits, it is better explained as an interpretive gloss for the longer reading (arising from an epexegetical reading of the genitive τοῦ εὐαγγελίου), and in terms of its testimony, the fact that it is both singular and corrected to the longer reading speaks against its originality. More importantly for our purposes, the temporal distance and textual dissimilarity between $\mathfrak{P}^{46}$ and GA 636 makes it highly unlikely that $\mathfrak{P}^{46}$ ever had this reading. In light of these factors, the sequence $\mathscr{T}_3$ should be ruled out from consideration, or the sequences should be assigned different prior probabilities in the calculations above to weigh them appropriately in light of external considerations.

More generally, this work makes the fundamental assumption that we can reduce the candidate texts for a lacuna to a few possibilities. In rich traditions like that of the New Testament, this is not an unreasonable assumption, as we usually find at least one reading in a variant passage that looks plausibly authorial. When this situation does not hold, if one can supply a conjectural reading with exceptional confidence, then a reconstruction using this conjecture can be supplied to this method. But again, the method will only have meaningful results if the conjecture effects a large change in the length of the reconstructed text.

5 Refinements

5.1 Accounting for subvariations

The more extensive a lacuna is, the more textual variants it is likely to contain. Even if the texts under consideration are expected to make up the most likely reconstructions, we might want to refine our estimates of their likelihoods based on the possibilities of errors, also lost in the lacuna, that would have affected the content of the lacunose text. In fact, we can do this, as long as we can assign probabilities to the errors.

To return to our example of Ephesians 6:19 in $\mathfrak{P}^{46}$, the text in the lacuna at the end of fol. 80$^{\text{v}}$ might have been affected by scribal errors known from elsewhere in the manuscript tradition. Two of these are the omission of τοῦ before στόματος (attested in GA 2492) and the omission of the μου after it (found in GA 0278). Since both of these omissions are singular and neither of their witnesses is known to have a close relationship with $\mathfrak{P}^{46}$, either omission would have to be an independent error on the part of the scribe of $\mathfrak{P}^{46}$. So, we now have four versions of each of our sequences $\mathscr{T}_1$ and $\mathscr{T}_2$ to consider: (1) the sequence without any errors; (2) the sequence with τοῦ omitted; (3) the sequence with μου omitted; and (4) the sequence with both τοῦ and μου omitted. The probabilities of the sequences without errors were already calculated in Section 3.3. The probabilities of the remaining sequences, calculated with the DP algorithm, are as follows:

$$
\begin{aligned}
&\Pr[\mathscr{T}_1 - \{\text{του}\}, \mathrm{L}=3] \approx 0.0007, \\
&\Pr[\mathscr{T}_1 - \{\text{μου}\}, \mathrm{L}=3] \approx 0.0006, \\
&\Pr[\mathscr{T}_1 - \{\text{του, μου}\}, \mathrm{L}=3] \approx 0.0023, \\
&\Pr[\mathscr{T}_2 - \{\text{του}\}, \mathrm{L}=3] \approx 0.0120, \\
&\Pr[\mathscr{T}_2 - \{\text{μου}\}, \mathrm{L}=3] \approx 0.0117, \\
&\Pr[\mathscr{T}_2 - \{\text{του, μου}\}, \mathrm{L}=3] \approx 0.0093.
\end{aligned}
$$

But now, these sequences are assumed to occur conditioned on the events of erroneous omissions in $\mathfrak{P}^{46}$, and the full-length sequences $\mathscr{T}_1$ and $\mathscr{T}_2$ are conditioned on the event that no such errors occurred. What are the probabilities of these events? Thanks to the work of Royse, we know that $\mathfrak{P}^{46}$ features '128 singular omissions … of one word' (Royse, 2008, p. 270) And thanks to Ebojo's tabulations, we know that the total number of lines (including reconstructed ones) for all the epistles found in $\mathfrak{P}^{46}$ is 4,822. Putting these together, we have a frequency of $q = 128/4822 \approx 0.0265$ small omissions per line. Assuming such errors are independent, the chance of both omissions in question occurring in $\mathfrak{P}^{46}$ is $q^2 \approx 0.0007$. The probability of only one occurring is $q(1-q) = (1-q)q \approx 0.0258$ for each error. The probability of neither occurring is $(1-q)^2 \approx 0.9476$. Let us denote by $\tilde{\mathscr{T}}_1$ (respectively $\tilde{\mathscr{T}}_2$) the union of the sequence $\mathscr{T}_1$ (respectively $\mathscr{T}_2$)

and its subvariations. The total probability that $\mathfrak{P}^{46}$ had $\mathscr{T}_1$ or any of its subvariations is therefore

$$\begin{aligned}\Pr[\tilde{\mathscr{T}}_1, L=3] &= (1-q)^2\Pr[\mathscr{T}_1, L=3] \\ &+ q(1-q)\Pr[\mathscr{T}_1 - \{\tau\text{ου}\}, \mathrm{L}=3] \\ &+ (1-q)q\Pr[\mathscr{T}_1 - \{\mu\text{ου}\}, \mathrm{L}=3] \\ &+ q^2\Pr[\mathscr{T}_1 - \{\tau\text{ου}, \mu\text{ου}\}, \mathrm{L}=3] \\ &\approx (0.9476)(0.0001) + (0.0258)(0.0007) \\ &+ (0.0258)(0.0006) + (0.0007)(0.0023) \\ &\approx 0.0002,\end{aligned}$$

and the probability that $\mathfrak{P}^{46}$ had the sequence $\mathscr{T}_2$ or any of its subvariations is

$$\begin{aligned}\Pr[\tilde{\mathscr{T}}_2, L=3] &= (1-q)^2\Pr[\mathscr{T}_2, L=3] \\ &+ q(1-q)\Pr[\mathscr{T}_2 - \{\tau\text{ου}\}, \mathrm{L}=3] \\ &+ (1-q)q\Pr[\mathscr{T}_2 - \{\mu\text{ου}\}, \mathrm{L}=3] \\ &+ q^2\Pr[\mathscr{T}_2 - \{\tau\text{ου}, \mu\text{ου}\}, \mathrm{L}=3] \\ &\approx (0.9476)(0.0103) + (0.0258)(0.0120) \\ &+ (0.0258)(0.0117) + (0.0007)(0.0093) \\ &\approx 0.0104.\end{aligned}$$

The hypothesis that $\mathfrak{P}^{46}$ had the shorter reading remains fifty-two times more likely than the hypothesis that it had the longer reading. Assuming that $\mathscr{T}_1, \mathscr{T}_2$, and their subvariations are the only possible reconstructions of the lacuna in $\mathfrak{P}^{46}$, the relative probabilities of the longer and shorter reading in $\mathfrak{P}^{46}$ become

$$\begin{aligned}\Pr[\text{reading }1 \mid \tilde{\mathscr{T}}_1 \text{ or } \tilde{\mathscr{T}}_2] &= \frac{\Pr[\tilde{\mathscr{T}}_1, L=3]}{\Pr[\tilde{\mathscr{T}}_1, L=3] + \Pr[\tilde{\mathscr{T}}_2, L=3]} \\ &\approx \frac{0.0002}{0.0002+0.0104} \\ &\approx 0.0148\end{aligned}$$

and

$$\begin{aligned}\Pr[\text{reading }2 \mid \tilde{\mathscr{T}}_1 \text{ or } \tilde{\mathscr{T}}_2] &= \frac{\Pr[\tilde{\mathscr{T}}_2, L=3]}{\Pr[\tilde{\mathscr{T}}_1, L=3] + \Pr[\tilde{\mathscr{T}}_2, L=3]} \\ &\approx \frac{0.0104}{0.0002+0.0104} \\ &\approx 0.9852,\end{aligned}$$

respectively.

5.2 Other linguistic assumptions

The probabilistic model for line length distributions from Section 2.1 and the DP algorithm from Section 2.2 are language-agnostic and can be used for any text, as long as the line lengths can be modeled as independent variables. Any probability distribution can be used in place of the normal distribution if it is more appropriate, but the assumption of independence is still required.

This method is also robust to situations where word-breaking conventions are known to be more lenient or stricter. For example, if a scribe is known to break lines at any point in a word without regard to syllabic boundaries, then this can be captured by tokenizing every letter in the sequence: in this way, the text υπερ εμου would be tokenized as υ-π-ε-ρ ε-μ-ο-υ. Alternatively, in languages or texts where breaking words across a line is uncommon, such as the Hebrew Bible (Ulrich, 1999, pp. 125–26; Beit-Arié, 2003, pp. 37–40), texts would be tokenized at the level of words, with no hyphens in the token sequence.

5.3 Finding the most likely segmentation

We can also modify the algorithm described in Section 2.2 to find the single most likely reconstruction of a text consisting of n tokens over L lines. This modified algorithm is in fact the standard segmentation algorithm, and it relies on the same DP technique. The only difference is that each entry in the DP table stores the maximum probability among all paths that end at a given node rather than the sum of all their probabilities. For the sequence $\mathscr{T}_1$ with the reading τὸ μυστήριον τοῦ εὐαγγελίου, the maximum-probability reconstruction of the three lacunose lines turns out to be Ebojo's reconstruction (with a probability of around 1.9446×10^{-5}), while for the sequence $\mathscr{T}_2$ with τὸ μυστήριον, it turns out to be Peterson's (with a probability of around 0.0022)—a testament to both scholars' transcriptional intuition.

6 Conclusion

From the application of the method described in this study, we can conclude with that $\mathfrak{P}^{46}$ was over fifty times more likely to have read τὸ μυστήριον than to have read τὸ μυστήριον τοῦ εὐαγγελίου in Ephesians 6:19. Moreover, if we assume that these two readings are the only readings $\mathfrak{P}^{46}$ could have had, then under these assumptions and the assumptions of our probabilistic model, we can say that $\mathfrak{P}^{46}$ had the shorter reading with probability near 99%.

This has several implications for our weighing of the evidence at this point of textual variation. First, the judgment of Kenyon, Hoskier, Zuntz, and Comfort on the reading of $\mathfrak{P}^{46}$ should be commended, and critical editions should print $\mathfrak{P}^{46\ \text{vid}}$ in favor of the shorter reading here. Second, Metzger's devaluation of GA 03's testimony on the basis of Western contamination, while it may be justified elsewhere in the Pauline corpus, must now be questioned here, as the alignment of

𝔓[46] and GA 03 opens the possibility that the shorter reading was simply transmitted normally from an early time. Indeed, the joint testimony of 𝔓[46] and 03 with the 'Western' witnesses is a force to be reckoned with in the manuscript evidence. Any argument that τὸ μυστήριον τοῦ εὐαγγελίου belongs in the authorial text of Ephesians 6:19 therefore should not rest on the manuscript evidence alone; a more comprehensive judgment of the readings on internal grounds is necessary.

For readers interested in experimenting with the approach outlined in this work or incorporating it into their own research, I have implemented the algorithm described here in the `calclac` Python script, which is accessible at https://github.com/jjmccollum/calclac. The code is open-source and available for free use and adaptation under the MIT License.

Funding

This work was supported by an Australian Government Research Training Program (RTP) Scholarship.

Conflict of interest

The author has no conflicts of interest to report.

Acknowledgements

I wish to thank my PhD supervisor Stephen C. Carlson for the questions that inspired this research and the suggestions that improved this paper, as well as the anonymous referees for feedback that made the argumentation clearer.

References

Beit-Arié, M. (2003). *Unveiled Faces of Medieval Hebrew Books: The Evolution of Manuscript Production – Progression or Regression?* Jerusalem: Magnes Press.

Bellman, R. (1961). On the approximation of curves by line segments using dynamic programming. *Journal of the ACM*, **4** (6): 284.

Best, E. (1998). *Ephesians. International Critical Commentary*. Edinburgh: T&T Clark.

Bruce, F. F. (2012). *The Epistle to the Ephesians: A Verse-by-Verse Exposition*. London: Pickering & Inglis. Bath: Creative Communications.

Comfort, P. W. (2015). *A Commentary on the Manuscripts and Text of the New Testament*. Grand Rapids, MI: Kregel Academic.

Comfort, P. W. and Barrett, D. P. (2019). *The Text of the Earliest New Testament Greek Manuscripts*. 3rd ed. Vol. **1**. Grand Rapids, MI: Kregel Academic.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2009). *Introduction to Algorithms*. 3rd ed. Cambridge, MA: MIT Press.

Dubois, J. -D. (1974). Ephesians IV 15: 'ἀληθεύοντες δὲ' or 'ἀλήθειαν δὲ ποιοῦν-τες': on the use of coptic versions for new testament textual criticism. *Novum Testamentum*, **16** (1): 30–34.

Duff, J. (1998). 𝔓[46] and the pastorals: a misleading consensus? *New Testament Studies*, **44**(4): 578–90.

Eadie, J. (1861). *A Commentary on the Greek Text of the Epistle of Paul to the Ephesians*. 2nd edn. New York, NY: Robert Carter.

Ebojo, E. B. (2014). *A Scribe and His Manuscript: An Investigation into the Scribal Habits of Papyrus 46 (P. Chester Beatty II – P. Mich. Inv. 6238)*. Ph.D. dissertation, University of Birmingham.

Ellicott, C. J. (1884). *St. Paul's Epistle to the Ephesians: with a Critical and Grammatical Commentary and a Revised Translation*. 5th edn. London: Longmans, Green & Co.

Hoehner, H. W. (2002). *Ephesians: An Exegetical Commentary*. Grand Rapids, MI: Baker Academic.

Metzger, B. M. (1994). *A Textual Commentary on the Greek New Testament*. 2nd edn. Stuttgart: Deutsche Bibelgesellschaft.

Peterson, J. W. (2020). *GA 1739: A Monk, his Manuscript and the Text of Paul's Letters*. PhD dissertation, University of Edinburgh.

Royse, J. R. (2008). *Scribal Habits in Early Greek New Testament Papyri*. New Testament Tools, Studies and Documents 36. Leiden: Brill.

Sanders, H. A. (1935). *A Third-Century Papyrus Codex of the Epistles of Paul*. University of Michigan Studies, Humanistic Series 38. Ann Arbor, MI: University of Michigan Press.

Schnackenburg, R. (1991). *The Epistle to the Ephesians: A Commentary*. Translated by H. Heron. Edinburgh: T&T Clark.

Ulrich, E. (1999). *The Dead Sea Scrolls and the Origins of the Bible*. Leiden: Brill.

Urdan, T. C. (2010). *Statistics in Plain English*. 3rd edn. New York, NY: Routledge.

Weiss, B. (1896). *Textkritik der Paulinischen Briefe. Texte und Untersuchungen zur Geschichte der altchristlichen Literatur 14.3*. Leipzig: J. C. Heinrichs.

Westcott, B. F. and Hort, F. J. A. (1881–82). *The New Testament in the Original Greek. 2 vols*. New York, NY: Harper.

Zuntz, G. (1953). *The Text of the Epistles: A Disquisition upon the Corpus Paulinum*. London and Eugene, OR: The British Academy/Wipf & Stock.

Notes

1. See, for a preliminary study of 𝔓[46], its scribal features, and its text, Sanders (1935); for its text-critical value, Zuntz (1953, pp. 14–57); for a recent and more thorough investigation into the habits of its scribe, as well as a new transcription, Ebojo (2014); and, for an even more recent transcription, the appendix of Peterson (2020).
2. A helpful introduction to these concepts that avoids overtechnicality can be found in Urdan (2010).
3. For the purposes of this study, a line is insufficiently extant if more than half of its letters must be reconstructed from lacunae.

4. This possibility is opened by Duff (1998), who observes an increasing character-per-page count towards the end of the papyrus, although Ebojo argues that this observation is due more to a gradual increase in the number of lines per page than to other factors (2014, pp. 204–34).

5. The assumption of independence is natural in this setting. If the scribe is assumed to have composed lines based on the constraints of line-breaking conventions and the objective of achieving a consistent line length, then the influence of one line's length on that of another should be negligible, if it exists at all.