2011

Examining a Misapplication of Nearest Shrucnken Centroid Classification to Investigate Book of Mormon Authorship

Paul J. Fields

G. Bruce Schaalje

Matthew Roper

Follow this and additional works at: https://scholarsarchive.byu.edu/msr

BYU ScholarsArchive Citation
Available at: https://scholarsarchive.byu.edu/msr/vol23/iss1/8

This Response to Criticisms is brought to you for free and open access by the All Journals at BYU ScholarsArchive. It has been accepted for inclusion in Review of Books on the Book of Mormon 1989–2011 by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Examining a Misapplication of Nearest Shrunken Centroid Classification to Investigate Book of Mormon Authorship</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Paul J. Fields, G. Bruce Schaalje, and Matthew Roper</td>
</tr>
<tr>
<td><strong>ISSN</strong></td>
<td>2156-8022 (print), 2156-8030 (online)</td>
</tr>
<tr>
<td><strong>Abstract</strong></td>
<td>Review of “Reassessing Authorship of the Book of Mormon Using Delta and Nearest Shrunken Centroid Classification” (2008), by Matthew L. Jockers, Daniela M. Witten, and Craig S. Criddle.</td>
</tr>
</tbody>
</table>
Examining a Misapplication of Nearest Shrunken Centroid Classification to Investigate Book of Mormon Authorship


Paul J. Fields, G. Bruce Schaalje, and Matthew Roper

Editor’s note: The above-referenced essay by Jockers, Witten, and Criddle (hereafter Criddle and associates) was answered by G. Bruce Schaalje, Paul J. Fields, Matthew Roper, and Gregory L. Snow in a technical paper entitled “Extended nearest shrunken centroid classification: A new method for open-set authorship attribution of texts of varying sizes,” Literary and Linguistic Computing 26/1 (2011): 71–88. We have invited Fields, Schaalje, and Roper to provide both a popularization of this important essay and a brief history of efforts to use what is called stylometry to identify the authors of disputed texts. In addition, because Professor Criddle has been involved in efforts to resuscitate the Spalding-Rigdon theory of Book of Mormon authorship, Roper and Fields were also invited to comment on that rather moribund explanation in a separate essay that immediately follows this one.

In 1834 the first anti-Mormon book was published in Ohio by E. D. Howe. Relying on testimony claimed to have been gathered by D. P. Hurlbut, a disgruntled former member of the Church of Jesus Christ of Latter-day Saints and sworn enemy to Joseph Smith, Howe argued that the Book of Mormon was based on an unpublished fictional tale by an unsuccessful amateur novelist, Solomon Spalding. Spalding lived in Conneaut, Ohio, between 1809 and 1812. Howe claimed that Sidney Rigdon somehow acquired Spalding’s unpublished manuscript and added religious material, thereby concocting the Book of Mormon.¹

The 1884 recovery of an original Spalding manuscript bearing little resemblance to the Book of Mormon led most critics to abandon the Spalding-Rigdon conspiracy theory.² This manuscript is


². Most Latter-day Saint and non-Latter-day Saint students of the issue have concluded that even if “Manuscript Story” was not the
known today variously as “Manuscript Story” or the “Oberlin manuscript.” Today, among those who reject Joseph Smith’s explanation of the Book of Mormon, a majority see Joseph Smith alone as responsible for the text and believe that the Spalding theory sheds no light on Book of Mormon origins. A minority of these critics continue to argue that the Book of Mormon was based on a hypothesized second or third, now-lost Spalding manuscript, though even the existence of such a manuscript has never been proved.3

A recent article by three Stanford researchers—Matthew Jockers, Daniela Witten, and Craig Criddle—is the latest in a series of stylometric investigations of Book of Mormon authorship.4 The Criddle and associates study applies a statistical methodology developed for genomics research,5 known as Nearest Shrunken Centroid (NSC) classification, to the question of Book of Mormon authorship. In contrast to previous wordprint studies, Criddle’s team concluded that the majority of the chapters in the Book of Mormon were written by either Solomon Spalding or Sidney Rigdon: “The NSC results are consistent with the Spalding-Rigdon theory of authorship,” and “our findings are consistent with historical scholarship indicating a central role for Rigdon in securing and modifying a now-missing Spalding manuscript” (p. 482). Although they claim to have discovered evidence for smaller contributions from Parley P. Pratt and Oliver Cowdery, the authors “find strong support for the Spalding-Rigdon theory of authorship. In all the data, we find Rigdon as a unifying force. His signal dominates the book, and where other candidates are more probable, Rigdon is hiding in the shadows” (p. 483).

We here examine the stylometric analysis presented by Criddle and associates. We first review past attempts—stylometric and otherwise—to analyze Joseph’s writing style. We review the strengths and weaknesses of those attempts and assess past authors’ success in meeting objections to their findings. We then address the validity of Criddle and associates’ methodology, its utility in dealing with questions of authorship in general, and its application to authorship of the Book of Mormon in particular. Lastly, we present the findings of our study extending the NSC methodology, which shows the naiveté and invalidity of Criddle and associates’ efforts to add a mathematical patina to an untenable historical hypothesis that has been long abandoned by virtually all serious scholars, whether believers or skeptics.

Prelude to Stylometry: Joseph Smith’s Writing Style

In 1976 Elinore Partridge performed a study of the characteristics of Joseph Smith’s writing style. She also studied the writings of several of his closest associates—Sidney Rigdon, Frederick G. Williams,
Parley P. Pratt, Oliver Cowdery, and Willard Richards. Partridge detected a characteristic tone in the Prophet’s writings.

In contrast to the dark visions of Calvinism and the dry, rational theology of Unitarianism, Joseph Smith’s pronouncements emphasize the wonder of existence and the love of humanity. Likewise, in contrast to the threats of wrath, judgment, and damnation, which one can find in the statements of some of the early church leaders, there is an undercurrent of understanding and compassion in those of Joseph Smith. Moments of discouragement and anger do occur; however, even at times when he laments the state of mankind, he tempers the observations with trust in God, love for his family, and hope for the future. The love of others, the pleasure in variety, and the joy in living which is apparent in the language of Joseph Smith give us some real sense, I believe, of what he must have been like as a leader and a friend.6

Partridge also found significant “markers” of Joseph Smith’s style that distinguish his writing from that of other Latter-day Saint leaders of his day. These include a tendency to form a structure of “interconnected sentences joined, like links in a chain, by simple conjunctions,” a characteristic that she found could often be detected even after Joseph’s work had been edited by others.7

Joseph Smith’s writing is characteristically marked by a series of related ideas joined by simple conjunctions: and, but, for. In his handwritten manuscripts, he used neither punctuation nor capitalization as sentence markers. When his writing has been edited, or when someone else wrote words which he dictated, the result is an unusually large number of sentences beginning with for, and, or but (almost three out of five sentences). On the other hand, Sidney Rigdon seldom used these conjunctions, and almost never used them at the beginning of sentences; on the average, only about one in twenty sentences begins with and, for, but. Rigdon’s sentences frequently begin with participial or prepositional phrases; for example, ‘Having shown . . .’ ‘From the foregoing we learn . . .’ which is a structure Joseph Smith seldom used. Sidney Rigdon regularly used phrases such as ‘in order that,’ ‘so that,’ or ‘the fact that,’ to introduce and link ideas. Joseph Smith almost invariably uses ‘that’ or ‘this’ instead. Joseph Smith’s images and examples are concrete, specific, and well-detailed, while Sidney Rigdon’s tend to be abstract and generalized.8

Partridge also noted Joseph Smith’s use of “pronouns and demonstratives which require specific referents” and the use of a “series of modifying phrases which must be attached to other words,” features that she notes “suggest a personality used to seeing things as an interconnected whole rather than as separate parts.”9 She saw this as evidence that “Joseph Smith is more comfortable with the spoken than with the written language. The long interrelated sentences, with no clear stopping point, are typical of an oral style. The

---

occasional repetitions or awkward constructions also indicate that he is writing as he speaks.”

Interestingly, Partridge also detected evidence that some elements of Joseph Smith’s style could be found even in works that he oversaw or directed others to write for him.

Joseph Smith’s influence can be seen in many of the works which he did not actually write himself. For example, I see signs of his collaboration in the *Lectures on Faith.* The sermons and discourses published in the *Times & Seasons* and parts of the *History of the Church* have obviously been well polished and heavily edited; however, in details and in general structures of the sentences, it is possible to identify characteristics of Joseph Smith’s style. Even when a scribe has obviously altered sentence structure to conform to a more standard, written style (that is, with definite marks of punctuation, capitalization, and clearer divisions between sentences), the interrelationships and internal references characteristic of Joseph Smith’s style remain. Occasionally, there are certain images and examples which indicate that a reported version of a sermon or speech has managed to capture the essential ideas and illustrations of Joseph Smith, although the language may have been dramatically altered.

Partridge’s findings suggest that there are distinct and significant differences between the writing styles of Joseph Smith and Sidney Rigdon. Consequently, analyses of the writing styles exhibited in the text of the Book of Mormon might provide insights into the question of the book’s authorship and particularly into Rigdon’s alleged role in its origin.

## Stylometry and the Book of Mormon

Stylometry uses statistical techniques to quantitatively describe the characteristics of an author’s writing style. It is based on the fundamental premise that authors write with distinctive word-use habits. For example, one commonly used method measures the frequency with which an author uses or does not use certain words or groupings of words. Identifying the word-use patterns in a text of unknown or questioned authorship and then comparing those patterns with the patterns in texts of known authorship can provide supporting evidence for or against an assertion of authorship. Although the proper term for this type of analysis is *stylometry,* the term *word-print analysis* is also sometimes used (in a loose comparison to fingerprint analysis). However, an author’s writing style is not nearly as precise, distinctive, unalterable, or unchanging as his or her fingerprints, and so the latter term is a potentially misleading exaggeration.

Over the last thirty years, researchers have conducted five major and several minor stylometric studies of the Book of Mormon. We will describe the major studies by Larsen et al., Hilton, Holmes, Criddle et al., and Schaalje et al.

### First Study: Word-Frequency Analysis

In 1980 Wayne Larsen, Alvin Rencher, and Tim Layton examined word frequencies in a precedent-setting analysis of the Book of Mormon.

---


indicators of writing style, they used noncontextual words—the words that play a grammatical role in forming the structure of a message but do not provide the information content of the message. Examples are a, an, but, however, the, to, with, and without. These words are also called function words since by themselves they do not convey the author's message but, rather, provide the framework for the author's message. Studying the function words in a text can indicate an author's personal manner of expressing his or her ideas since they do not indicate what the author says but the way he or she says it.

The Larsen et al. researchers used three statistical techniques—Multivariate Analysis of Variance (MANOVA), Cluster Analysis (CA), and Linear Discriminant Analysis (LDA)—to test for differences in the frequencies of noncontextual words. MANOVA is a method of testing for homogeneity (degree of similarity) within groups of items. CA is a method that can identify which items are closest to each other among all items compared. LDA is a method for determining a set of mathematical functions (discriminant functions) that can be used to classify items into categories based on their characteristics. The three methods produced consistently congruent results, which are highlighted below using LDA to summarize the findings.

In stylometric analysis, LDA can compare the word-frequency profile in a block of text to the profile of each candidate author and then assign that block of text to the author with the most similar style. It does this by measuring how closely the word profile in the test block matches the average word profile of each author. A plot of the test texts using the discriminant functions as the axes of the graph can display how well the texts correspond to each author.

In the Larsen et al. study, the researchers segmented the entire text of the Book of Mormon into 2,000-word text blocks aligned with each of the twenty-one purported authors in the book. Then they tested whether there was evidence that the text blocks displayed a consistent style across the blocks (indicative of one author for all the texts) or whether there was evidence of differing styles (congruent with the claim that the Book of Mormon texts came from different writers).

For comparison they also examined texts from Joseph Smith, Sidney Rigdon, and Solomon Spalding, along with texts from Oliver Cowdery, Parley P. Pratt, and W. W. Phelps, all of whom they referred to collectively as “nineteenth-century authors.”

Larsen's team showed that the text blocks from the Book of Mormon were consistently classified as separate from those of the nineteenth-century authors. This is shown in figure 1. Further, they showed that each Book of Mormon author is consistently similar to himself but consistently different from the other authors. This is illustrated in figure 2, which shows the texts grouped into separate clusters by author. For simplicity in illustrating the results, figure 2 shows the clusters for only Nephi, Alma, and Mormon—the three major authors in the Book of Mormon. Joseph Smith's cluster is also shown in figure 2, and it stands apart from the Nephi, Alma, and Mormon clusters.

After repeatedly analyzing all the texts and all the candidate authors, Larsen's team found the statistical evidence of differences between the writings of the purported authors to be striking. They concluded that “distinct authorship styles can be readily distinguished within the Book of Mormon...
Mormon, and the nineteenth-century authors do not resemble Book of Mormon authors in style."¹³

D. James Croft, a statistician at the University of Utah, raised several questions in critique of the Larsen et al. analysis.¹⁴ In essence he asked the following:

1. Is the basic assumption of stylometrics—that authors’ writing styles can be characterized by measurable features—valid?
2. Does the modern Book of Mormon edition used by Larsen et al. exhibit the same stylistic patterns as those in the original 1830 edition?
3. Was the phrase “and it came to pass that” recognized by Larsen et al. as a possible indicator of content differences rather than author differences?
4. Were the results of the analysis due to style differences among the purported authors or to topic differences among the texts?

When Croft’s review of the Larsen et al. study was published, it was accompanied by a well-reasoned reply by the researchers to all the issues he raised.¹⁵ We offer here some additional analysis in further rebuttal.

Croft’s first point—the validity of stylometry—has been answered by the continuing and increasingly successful use of stylometric methods similar to those employed by Larsen’s team. The methodology has been validated repeatedly and is a well-accepted analytical approach in

---

literary analysis. Even other critics of the Larsen et al. study, such as David Holmes,\(^{16}\) do not dispute the validity of the methodology. However, most stylometry practitioners would agree with Croft that the methodology has limitations and that it is only as valid as the stylistic measures used in the analysis.

Croft’s second point—the use of a modern edition of the Book of Mormon—turns out to be a nonissue when we examine the effects of editorial changes to the book. Figure 3 overlays plots of word-use frequencies from sequential 5,000-word blocks of both the 1830 and 1980 editions of the Book of Mormon. The editorial changes to the Book of Mormon over 150 years appear to have been made nearly proportional throughout the book since the patterns present in one edition are almost exactly reproduced in the other. For the main purpose of the Larsen et al. study, it did not matter which edition the researchers used.

Croft’s third point—the possible effect of the frequently occurring phrase “and it came to pass”—is insightful. However, subsequent studies we have conducted showed no detectable differences in the results of stylometric analyses that include the words in the phrase “and it came to pass” as separate words, treat the phrase as one word, or delete those words entirely from the frequency counts when they occur in that phrase.

Croft’s fourth point—results due to style or topic differences—is well-taken. The consistent difference between writings attributed to Mormon and those attributed to Nephi or Alma could be due to content differences instead of stylistic differences, since Mormon’s writings tend to be historical narrative while Nephi’s and Alma’s writings tend to be doctrinal discourse. However, there can be little question that the Larsen et al. study showed, at a minimum, that the texts purported to be written by Nephi and Alma exhibit internally consistent but highly distinct authorship styles as measured by their use of noncontextual words, even though both authors were discussing the same topics.

Other criticisms of the Larsen et al. study have come forward more recently. The 2008 paper by Criddle and associates questioned the Larsen et al. approach of grouping verses and partial verses into blocks of words “based on their understanding of speakers (or characters) in the Book of Mormon” (p. 467). However, this criticism is misguided since such grouping was appropriate given that Larsen’s team was testing a hypothesis of multiple authorship. A second point raised by Criddle and associates was that even if the texts were carefully grouped, they might be “composites containing different fractional contributions from different nineteenth-century authors” (p. 467). Although this could be true, the

---

Fig. 3. Comparison of word-use frequencies in the 1830 and 1980 editions of the Book of Mormon. Texts from the 1830 edition are shown as green dots, and texts from the 1980 edition are shown as blue dots. For every green dot, there is a corresponding blue dot in close proximity for the same block of text, indicating close similarity between the two editions.

---

consistent clustering of writings due to purported Book of Mormon speakers would imply a remarkable compositing process, a process in which the different nineteenth-century authors contributed consistent but different proportions of text for each of the purported authors. Finally, Criddle and associates state that biblical-sounding words such as behold, forth, lest, nay, O, unto, wherefore, and yea might account for observed differences between Book of Mormon text blocks and the text blocks of the nineteenth-century authors in the study. However, the Larsen et al. study did not use those words, so perhaps Criddle and associates misread the word lists used by Larsen et al. We discuss in detail the paper by Criddle and associates later in this article.

On the whole, even after the thoughtful criticism of the Larsen et al. study is accounted for, the results of that early study continue to provide persuasive support for the claim that the Book of Mormon is the work of multiple authors and not the work of any of the likely nineteenth-century candidates.

Second Study: Word-Pattern Ratios Analysis

In a subsequent study, John Hilton took a different approach to studying stylistic patterns in the Book of Mormon.17 Intrigued but uncertain of the Larsen et al. results, Hilton set out to see if he could replicate their results using a study designed to accommodate Croft’s criticisms. Rather than noncontextual word frequencies as in Larsen et al., Hilton used “noncontextual word-pattern ratios.” Word-pattern ratios measure the rates an author uses words in four categories:

1. Specific words in key positions of sentences (e.g., the as the first word of a sentence)
2. Specific words adjacent to certain parts of speech (e.g., and followed by an adjective)
3. Collocations of words (e.g., and followed by the)
4. Proportionate pairs of words (e.g., no and not, all and any)

Hilton used the sixty-five word-pattern ratios developed by A. Q. Morton that he had shown to be useful in authorship studies for other religious texts as well as secular texts.18 One of the advantages of using word-pattern ratios is that the potentially problematic phrase “and it came to pass” can only partly affect one of the sixty-five word-pattern ratios, so its impact on the analysis was negligible in Hilton’s approach.

Using primarily the printer’s manuscript of the Book of Mormon, Hilton applied his procedure to 5,000-word blocks of text to ensure the reliability of the style measures since in larger text blocks an author’s writing habits and stylistic propensities should assert themselves more strongly than in smaller texts. He compared texts attributed only to Nephi or Alma to control for topic differences and then texts known to be authored by Joseph Smith, Oliver Cowdery, or Solomon Spalding. He compared each author to himself and then each author to every other author. The result demonstrated that the stylistic patterns in the Nephi, Alma, Smith, Cowdery, and Spalding texts were consistent within themselves but distinctly different from one another.

---


This evidence argues strongly for the assertion that the Nephi and Alma texts were written by different authors, and against the idea that Joseph Smith, Oliver Cowdery, or Solomon Spalding was the author of the Nephi or Alma texts. Hilton stated:

It is statistically indefensible to propose Joseph Smith or Oliver Cowdery or Solomon Spaulding as the author of the 30,000 words from the Book of Mormon manuscript texts attributed to Nephi and Alma. Additionally these two Book of Mormon writers have wordprints unique to themselves and measure statistically independent from each other in the same fashion that other uncontested authors do. Therefore, the Book of Mormon measures [as being] multiauthored, with authorship consistent to its own internal claims.19

Hilton’s results corroborated the Larsen et al. results even though Hilton used an entirely different technique.

Third Study: Vocabulary Richness Analysis

In 1992 David Holmes published the results of a stylometric analysis of the Book of Mormon using another approach, one he had developed as a doctoral student. He attempted to show that measures of “vocabulary richness” could be used for authorship attribution.20 Vocabulary richness measures attempt to quantify an author’s style based on his or her lexical variety in word choices. As stylistic features, Holmes computed a standardized measure of once-used words, a standardized measure of twice-used words, a measure of lexical repetitiveness, and two estimated parameters for a theoretical model of word frequencies in writing. The first three measures were calculated for the total vocabulary in the texts, while the last two were calculated for nouns only.

Holmes compiled fourteen 10,000-word blocks assigned to six Book of Mormon authors, combined sections 1 through 51 of the Doctrine and Covenants into three 10,000-word blocks, combined an assortment of writings by Joseph Smith into three 6,000-word blocks, included the Book of Abraham from the Pearl of Great Price as one text, and extracted three 12,000-word blocks from Isaiah. He then used Principal Components Analysis (PCA) to search for separations among the clusters of texts.

PCA takes a set of multidimensional points and projects them into two dimensions. As an analogy, imagine the outline of a three-dimensional object such as a pencil projected by an overhead projector onto a flat, two-dimensional screen. Its projected image could look like a dot or like an arrow, depending whether the pencil is oriented vertically or horizontally. The PCA procedure determines how to rotate a set of points so the greatest separation among the points can be seen. This is a useful way to visually explore the data in two dimensions for possible relationships among points in many dimensions. The first and second principal components define the two-dimensional space.

Using PCA applied to the vocabulary richness measures, Holmes found that the Joseph Smith texts clustered together, the Isaiah texts clustered together, and all but three of the other texts clustered together. Figure 4 presents a PCA plot of Holmes’s results. From this he concluded that the writings of Mormon, Lehi, Nephi, Jacob, and Moroni were not stylistically different.

---

20. See note 16.
Subsequent research has shown that Holmes’s vocabulary richness stylistic measures are weak discriminators of authorship. For example, when testing texts of undisputed authorship, correct classification rates were 96 percent using non-contextual word frequencies, 92 percent for non-contextual word-pattern ratios, but only 23 percent for vocabulary richness measures.21 In statistical terms, a method’s ability to find differences is called “power.” A weak discriminator, such as the vocabulary richness measure, can lack the power to find differences even when they are present.

When a method cannot find differences that are known to exist in the data, and then subsequently does not find a difference between two items, such a result is not convincing evidence that no true difference exists between those items.

Consequently, the correct interpretation of Holmes’s finding is not that “there are no differences among the tested authors,” but rather that he “found no evidence of difference.” Not finding evidence of difference may therefore say little about the subject of the test but can be an indication of the test’s inadequacy. This was the case for Holmes’s Book of Mormon study—he was using a technique with low power. Such a situation is analogous to using a low-powered microscope when a high-powered instrument is needed: his “instrument” was inadequate for the research he was attempting, leaving him unable to discern features that were, in fact, present.

Although in concept vocabulary richness analysis seems like it should be useful, in practice it has been shown to be unreliable. In fact, after his early work in stylometrics, Holmes subsequently discontinued the use of vocabulary richness measures and employed other techniques in his work. We conclude that the Holmes study serves only to show the limitations of vocabulary richness analysis, while providing no insight into the question of Book of Mormon authorship.

Fourth Study: Nearest Shrunken Centroid Analysis

Sixteen years after the Holmes study, Matthew Jockers, Daniela Witten, and Craig Criddle tried to take an innovative approach to authorship attribution by applying an analytical method developed for the classification of tumors in genomics research.22 The technique is called Nearest Shrunken Centroid (NSC) classification. It takes a set of items of known origin(s) and compares them to a set of items of unknown origin(s) by determining the distances between the centers.


(centroids) of the groups of items. The technique seeks to identify the centroids that are nearest to each other. "Shrinkage" is a statistical technique to combine all available information in a way that can reduce the uncertainty in estimating the distances between the centroids. The distance between the centroids is considered a surrogate for similarity. When centroids are relatively close to each other, this is taken to indicate relative similarity. Conversely, when the centroids are relatively far apart, this is taken to indicate relative dissimilarity. NSC calculates the probability of relative similarity.

When applied to stylometry, NSC develops a classification rule based on stylistic characteristics—such as word frequencies—in a set of texts with known authorship and then uses that classification rule to assign texts of questioned authorship to the author whose style is closest. The closer a test text of an unknown author is to the centroid of a known author’s texts, the greater the likelihood that the writing style exhibited in the test text matches the writing style of the known author. The analysis is complex since each word frequency is a dimension in which “distance” must be measured. If a researcher uses one hundred word frequencies, the analysis is a one-hundred-dimensional problem.

Criddle and associates applied NSC to the Book of Mormon in an attempt to find evidence in support of the Spalding-Rigdon theory. Their set of texts for candidate authors included Solomon Spalding, Sidney Rigdon, Parley P. Pratt, and Oliver Cowdery. They also included Isaiah and Malachi (combined as one author) as a positive control and Joel Barlow and Henry Wadsworth Longfellow (nineteenth-century authors) as negative controls. The texts varied widely in length from 114 to 17,797 words. Their test texts were the individual chapters of the Book of Mormon, which varied in length from 95 to 3,752 words.

As stylistic features, Criddle and associates used the relative frequencies of the most common 110 words in the Book of Mormon that were used at least once by each candidate author. Although their list contained mainly function words, they retained some lexical words as well. From their analysis they concluded that the evidence showed with high probability that Spalding and Rigdon were the principal authors of the Book of Mormon.

However, there were very significant problems with this study. We will discuss the following problems:

- Failing to include Joseph Smith as a candidate author
- Misapplying a closed-set technique for an open-set problem
- Confusing “closest” to mean “close”
- Misinterpreting relative probabilities as absolute probabilities
- Ignoring a high rate of false classifications
- Using circular statistical thinking
- Disregarding statistical problems of homogeneity and multiplicity
- Confounding the primary candidate author’s differing writing styles

Failing to Include Joseph Smith as a Candidate Author

Considering the lack of unanimity on the question of Book of Mormon authorship, even among critics, it is strikingly odd that Criddle and associates would choose to exclude Joseph Smith from the list of potential authors. A substantial majority of critics favor some version of the Joseph Smith composition theory, which sees Joseph Smith as the book’s author. Latter-day
Saints, on the other hand, who understand the Book of Mormon as divinely revealed scripture, acknowledge Joseph Smith as a human mediator of the revealed word (2 Nephi 31:3; D&C 1:24) and may be interested in the degree to which the Prophet’s language may have influenced the translation of the text. Consequently, the question of Joseph Smith’s influence on the text of the Book of Mormon is one of considerable interest to both Latter-day Saints and non–Latter-day Saint students of religion.

In an attempt to justify this significant omission, Criddle and associates noted that Joseph Smith usually dictated his writings to others. They cite Dean Jessee, the leading authority on Joseph Smith’s personal writings, who explains (like Partridge above) that Joseph Smith appears to have been much more comfortable as a speaker than a writer and that, consequently, the majority of his writings are not in his own hand but in that of scribes (p. 469). However, Criddle and associates make the astonishing assertion that even Joseph Smith’s holographic writings—those written in his own hand—are unreliable examples of Joseph’s written style. “In the case of Joseph Smith, we do not believe that even the small number of letters written in his own hand can be reasonably attributed to him. Moreover, were we to concede the reliability of these few letters, we would still not have enough text to constitute an ample sample of known authorship” (p. 486). The authors make two claims: (1) that the writings of Joseph Smith in his own hand are not a reliable source of data reflecting his writing style and (2) that there are not enough of these writings to utilize in a wordprint study. The first claim is mystifying, and the second claim is unjustified.

First, their hyperskepticism about Joseph Smith’s holographic writings is not supported by historians. Dean Jessee, whom they cite in support of this claim, states: “The real importance of Joseph Smith’s holographic writings (the writings he produced with his own hand) lies in their being his expression of his own thoughts and attitudes, his own contemplations and reflections. They not only reveal idiosyncrasies of his education and literary orientation but also clearly reflect his inner makeup and state of mind—his moods and feelings. Furthermore, they provide a framework for judging his religious claims.”

In a separate article, Jessee explains, “One of the best avenues, which is undistorted by clerical and editorial barriers” for studying Joseph Smith as a speaker and a writer, “is the Prophet’s holographic writings—those materials produced by his own hand and hence by his own mind.” Writing that captures an author’s “inner makeup and state of mind” is essential when performing a stylometric analysis.

The authors’ second claim—that even if one wanted to use holographic material from Joseph Smith there would not be enough to be useful—seems disingenuous, given that they use samples from other candidate authors with sizes as small as only 114 words (p. 471). Available holographic material potentially includes (1) holographic portions of Joseph Smith’s 1832 history (1,016 words); (2) portions of Smith’s 1832–34 Kirtland, Ohio, Journal (1,589 words); (3) portions of his autobiography. 23

---

25. “This document is the earliest extant attempt by the Prophet to write a history of his life, and his only autobiographical work containing his own handwriting.” Dean C. Jessee, ed., The Papers of Joseph Smith, Volume 1: Autobiographical and Historical Writings (Salt Lake City: Deseret Book, 1989), 1.
26. “Despite its brevity, this first journal contains more of JS’s
1835–36 Kirtland, Ohio, journal, which contains seven entries (four manuscript pages) in his own hand (529 words); 27 (4) three letters partly in the hand of Joseph Smith and partly in the hand of another writer (899 words); 28 and (5) twenty-four letters entirely written in Joseph Smith’s handwriting totaling over 12,039 words. 29 While these holographic texts are small in quantity when compared to the entire corpus of historical documents dictated by or prepared under the direction of Joseph Smith, it seems reasonable to expect that a serious researcher would use these materials and could thereby obtain a reliable and adequate sample for the purposes of authorship analysis.

After the paper by Criddle and associates was published and this most obvious error in their analysis was pointed out, Matthew Jockers attempted to justify the error in an unpublished manuscript. A review and analysis of that manuscript is provided in the appendix to this paper.

Misapplying a Closed-Set Technique for an Open-Set Problem

In their study, Criddle and associates treat the set of candidate authors as a “closed set,” assuming that they knew with certainty that the true author was one of the authors in their candidate set. Although such an assumption would be appropriate when using NSC in the genomic studies for which it was originally developed, this is not appropriate in most authorship attribution studies. The case of The Federalist Papers is a situation where the true author was known to be in the candidate set—the twelve disputed articles were written by either Alexander Hamilton or James Madison, and by no one else. Such a well-defined closed-set problem as The Federalist Papers is a rarity in authorship attribution studies.

Although Criddle and associates show that NSC performed well in an analysis of The Federalist Papers, this is to be expected when applying a closed-set procedure to a closed-set problem. The case of the Book of Mormon is clearly not the same type of problem. In their study, Criddle and associates did not allow for the possibility that the Book of Mormon was a translation of writings authored many centuries ago, nor (as discussed in the previous section) did they consider the option that most secular critics deem most plausible: that Joseph Smith himself was the
author. Not allowing for either possibility prejudiced their study’s results from the start.

To understand the consequences of naively applying the NSC classification technique indiscriminately, let’s consider four cases in which we use a closed set of candidate authors when clearly an open-set should be used.

First, if we naively apply NSC to *The Federalist Papers* using Criddle and associates’ set of candidate authors and using their way of interpreting the results, we find with 99 percent probability that Sidney Rigdon wrote thirty-four of *The Federalist Papers* published in 1788, before he was even born (he was born in 1793). If we ignore important potential authors, Criddle and associates’ technique will mislead us with a high level of confidence in a misattribution.

Next, if we propose that the Spalding-Rigdon theory applies to the King James Bible as well as to the Book of Mormon and then naively apply NSC to the Bible using Criddle and associates’ closed set of authors, we find with 99 percent probability that Sidney Rigdon wrote about 30 percent of the Bible. If one wishes to attach any validity to Criddle and associates’ finding about Rigdon as an author of the Book of Mormon, he or she must also attach the same level of validity to Rigdon’s authorship of the Bible.

Similarly, if we concoct the absurd scenario that one or more of a closed set of five early anti-Mormon writers—Alexander Campbell (1831), Eber D. Howe (1834), Daniel Kidder (1842), Tyler Parsons (1841), and Walter Scott (1841)—wrote the Book of Mormon, when we naively apply NSC as Criddle and associates did in their study, we find that Parsons was the principal author of the Book of Mormon since NSC attributed 65 percent of the chapters to him with greater than 99 percent probability.

Finally, applying that naive approach to the paper under review and using its candidate set of authors, we find with 99 percent probability that Oliver Cowdery (who died in 1850) wrote the Jockers, Witten, and Criddle paper published in 2008. Clearly, this approach produces absurd results when naively employed unless Criddle and associates are willing to disavow authorship of their own paper!

We can see from these examples how easily researchers could deceive themselves into thinking they had found evidence in support of a hypothesized authorship attribution regardless of how impossible or baseless it might be. We can illustrate this graphically with an additional seemingly plausible example. Let us propose that the Book of Mormon was written by either Solomon Spalding or James Fenimore Cooper, the author of Last of the Mohicans. We base this conjecture on the simple facts that both authors lived during the same time period (Spalding 1761–1816 and Cooper 1789–1851), both wrote their documents prior to the publication of the Book of Mormon in 1830 (Spalding’s “Manuscript Story” circa 1800 and Cooper’s Last of the Mohicans in 1826), both wrote in the same genre (historical fiction), both used the same geographic setting for their stories (northeastern North America), and, most importantly, both used Native Americans as their subject matter. Now, since the Spalding-Rigdon theory alleges that Spalding’s work was the basis for the historical narrative in the Book of Mormon with Rigdon contributing the doctrinal content, and since Rigdon is not included in our Spalding-Cooper theory, we test our theory by examining the writing styles in only the chapters of the Book of Mormon that are primarily historical in nature. We use the same noncontextual words as
Criddle and associates to determine the word-use frequencies in the texts. Figure 5 presents a principal components plot of the Book of Mormon texts along with Spalding’s “Manuscript Story” texts and Cooper’s Last of the Mohicans texts. We can easily see that the writing styles of nineteenth-century authors Spalding and Cooper are more similar to each other than to texts from the Book of Mormon. Spalding’s “Manuscript Story” is no more similar to the Book of Mormon than is Cooper’s Last of the Mohicans.

Criddle and associates assert that when any potpourri of authors is collected and then a closed-set procedure is used to assess attribution, the style of at least one of the candidate authors will always be identified as “closest to” the style of the author of the test text. It is also equally clear that “closest to” (a relative comparison) does not necessarily mean “close to” (an absolute comparison), and therefore caution is always necessary in interpreting the results.

Confusing “Closest” to Mean “Close”

The logic of Criddle and associates’ approach is no different than asking, “Choosing among Boston, New York, and Chicago, which city is closest to Los Angeles?” and then, upon finding that there is a 99 percent probability that Chicago is the closest, concluding that “Chicago is the city in the United States that is closest to Los Angeles.” In addition, finding that one city of three candidate cities is “closest” to some target city does not mean the cities are necessarily “close” to each other. Just as Chicago might be closest to Los Angeles given the closed set consisting of Chicago, New York, and Boston, certainly Chicago is not closest given the open set of Chicago, New York, Boston, or any other city in the United States.

Also, since Chicago and Los Angeles are half a continent apart, few people would say they are “close” to each other, let alone that they are the same city. In similar fashion, Criddle and associates assert that when, according to their calculations, the writing style in a test text is “closest” to one author’s style, then the two styles are “close.” In fact, they imply that the styles are close enough to be considered identical. This is nonsense.

Misinterpreting Relative Probabilities as Absolute Probabilities

Since the NSC technique is a closed-set analysis technique, the probabilities of closeness of writing style calculated by Criddle and associates can be interpreted only as relative probabilities. That is to say, the probabilities are only relative to the authors in the closed set of candidate authors. Yet Criddle and associates present their calculations
as absolute probabilities, which would require that all possible outcomes had been included in the computations. They obviously misinterpreted the probabilities as saying, for example, that there is a greater than 99 percent chance that Rigdon wrote a particular text, when the correct interpretation is that there is a greater than 99 percent chance that Rigdon’s writing style is closer to the style exhibited in a particular text compared to the other author candidates used in the analysis. They give the false impression that their 99 percent computation is an absolute measure applicable to all possible candidates, when it applies only to the specific set of authors they choose to include.

Ignoring a High Rate of False Classifications

In determining the reliability of an analytical technique, a researcher will use “positive controls” and “negative controls.” In a stylometric analysis, authors will be included in the candidate set who are known to have contributed some of the test texts. These serve as positive controls to test if the method can identify authors for whom some texts should be attributed. Conversely, authors will be included in the candidate set who are known not to have contributed any of the test texts. These serve as negative controls to determine whether the method can exclude authors to whom texts should not be attributed. In the Criddle and associates study, texts by Isaiah and Malachi were composited together into one set of texts to use as a positive control since those ancient prophets had definitely authored some of the chapters in the Book of Mormon. Similarly, texts by Joel Barlow and Henry Wadsworth Longfellow were included as negative controls since those poets had definitely not authored any of the chapters.

Although the NSC technique was able to exclude Barlow and Longfellow, it did poorly with Isaiah and Malachi. NSC correctly classified twenty of twenty-one Isaiah and Malachi chapters, but it misclassified forty-two other chapters as being authored by Isaiah and Malachi. A technique that makes twice as many false classifications as correct classifications for the control texts cannot possibly be considered to be a reliable technique. Consequently, whatever other classifications it produced must be viewed as unreliable and uninterpretable. It is astonishing that Criddle and associates ignored their technique’s high rate of false classifications.

Using Circular Statistical Thinking

Statistical methods are not foolproof and must be used correctly to produce reliable results. For example, if a statistical method is used to fit a straight line to a set of data for two variables, \( x \) and \( y \), the method will give a straight line even if the data follow a curved pattern (see fig. 6). To deal with this issue, an assessment of the data relative to a proposed model should be carried out before fitting the model, and a confirmatory goodness-of-fit test should be done after fitting. Concluding that a straight-line model is appropriate simply because a straight line can be fit to a data set is obviously fallacious circular reasoning. Justifying straight-line predictions of \( y \) from values of \( x \) by saying that the predictions are correct assuming the straight-line model to be correct could lead to grossly incorrect predictions.

Criddle and associates made exactly this kind of mistake in their stylometric analysis. Without checking the fit, they assumed that every chapter of the Book of Mormon was written by one of their seven candidate authors (Rigdon, Spalding, Cowdery, Pratt, Isaiah/Malachi, Longfellow, and
Barlow). Then, using NSC to assign each of the Book of Mormon chapters to one of their set of authors, they concluded that since almost all of the noncontrol chapters were assigned to one of the noncontrol authors, they had discovered “strong support for the Spalding-Rigdon theory of authorship.” However, they had simply forced a model on the data and then circularly concluded that agreement of the predictions with their model provided evidence for their model.

Just because a model can be mathematically fit to some set of data does not mean it is the right model for the data. Neglecting to check whether the data were actually consistent with the model applied to them is a serious mistake, whether due to ignorance, inexperience, or willful blindness. Such verification is not easy to do with highly multivariate data such as stylometric data, but it is nonetheless necessary if one wishes a reliable analysis with interpretable results.

Disregarding Statistical Problems of Homogeneity and Multiplicity

Criddle and associates disregarded two fundamental statistical issues in their analysis: homogeneity of variance and multiplicity in hypothesis testing. The NSC procedure employed by them assumes that the variance of the word frequencies in the text blocks is the same (homogenous) for all of the text blocks. However, the text blocks in their study ranged in size from about one hundred words to more than fifteen thousand words. The variances calculated in text blocks spanning such a wide range will produce widely differing variance estimates. Criddle and associates did not realize that the NSC results will have questionable reliability when the homogeneity of variance assumption is violated.

Further, the study simultaneously classified 239 chapters from the Book of Mormon into seven authorship categories in a single statistical procedure. In such situations of multiple simultaneous classifications (multiplicity), some of the calculated probabilities will appear to be unusually large even though they were simply the result of chance. The probability that a text should be associated with a certain candidate author versus another can be overstated. For example, a text on the stylistic fringe of an author’s cluster of texts can stray into a nearby author’s cluster and appear to be closer to that author rather than the true author. Consequently, classification probability results must be interpreted collectively rather than individually so as not to overinterpret the results. Criddle and associates did not account for the multiplicity effect of classifying a large set of test texts when interpreting their results.

---

Fig. 6. A straight-line model fit to curved data. Just because a model can be fit to the data does not mean that it is the appropriate representation of the data. A researcher cannot assume that the model fits the data. He or she must demonstrate the model’s appropriateness by a goodness-of-fit test.

30. Variance is a measure of the dispersion or inconsistency among multiple observations of the same phenomena.
Confounding the Primary Candidate Author’s Differing Writing Styles

Finally, Criddle and associates assumed that an author’s writing style is constant throughout his or her lifetime. They should have investigated this assumption, particularly for their prime candidate. If they had, they would have discovered that the Rigdon texts written prior to 1846 show evidence of being systematically different from those written after 1863. Most notably, while both sets of texts are distant from the Book of Mormon chapters, Rigdon’s post-1863 writings are closer to the Book of Mormon than his pre-1846 writings. This is illustrated in figure 7.

If Rigdon had been involved somehow in composing the Book of Mormon, the “early Rigdon” rather than the “late Rigdon” texts would be closer to the Book of Mormon chapters. The opposite is the case, and this clearly contradicts the Spalding-Rigdon theory for Book of Mormon authorship. The existence of two distinct Rigdon styles makes interpreting Criddle and associates’ results highly problematic. Since they composited texts containing Rigdon’s two differing writing styles, they blur what they call “Rigdon’s signal.” Whatever “signal” might have been present in the texts contained two styles rather than one distinguishable style. Because the Rigdon writing style closest to the Book of Mormon was chronologically disjointed from the book’s publication, whatever “signal” that Criddle and associates thought they had found was a signal that came into existence over thirty years too late to support their contention. (If anything, the results might suggest that Book of Mormon language and style influenced Rigdon’s later style, and not the other way around.)

Fifth Study: Extended Nearest Shrunken Centroid Analysis

Most recently, in a new study we developed a modification to the closed-set Nearest Shrunken Centroid classification method (NSC) to enable it to be applied to open-set classification problems—Extended Nearest Shrunken Centroid (ENSC) classification. The open-set modification allows for the existence of an unknown candidate author with a distribution of characteristic features nominally consistent with the test text and incorporates this possibility into the calculation of the probabilities that the writing styles are similar. Without including the possibility of an unknown author, if the candidate set does not include the true author (using a closed-set approach for an open-set situation), the calculated probabilities can be grossly overstated and lead to entirely erroneous interpretations. The ENSC technique also accounts for differences in text sizes and for the effect of multiple simultaneous comparisons.

Fig. 7. PCA plot of early and late Rigdon texts. The early Rigdon texts were written before 1846 and are shown as solid red dots. The late Rigdon texts were written after 1863 and are shown as open red dots. The two clusters do not overlap, suggesting strongly that Rigdon’s early writing style had evolved into another style later in his life.

In applying this technique, we used the same 110 characteristic words as in the Criddle and associates study, as well as their chapter-by-chapter designation of text blocks from the Book of Mormon. However, we used only the “early Rigdon” texts to represent Rigdon’s style and included Joseph Smith as a candidate author along with the possibility of an unknown author.

The open-set ENSC method produced far different results than those reported by Criddle and associates.

To illustrate the results, first we present in figure 8 plots showing the Book of Mormon texts along with the texts from the candidate authors. The 110 dimensions of the word frequencies have been projected into a two-dimensional space defined by the first two principal components.
The sequence of plots shows the clusters of texts for each author individually in relation to the Book of Mormon texts. Overall, it can be seen that in all cases the candidate authors’ texts cluster separately from the Book of Mormon texts. Further inspection shows that Solomon Spalding’s cluster is actually the farthest away from the Book of Mormon, with Sidney Rigdon’s cluster almost as far removed. Parley Pratt’s cluster is grouped with Spalding and Rigdon.

Next we can see in figure 8 that Oliver Cowdery’s and Joseph Smith’s clusters are closer to the Book of Mormon cluster than Spalding’s or Rigdon’s, thus confirming Criddle and associates’ error in not including Joseph Smith in their set of candidate authors. Finally, we note that, as a group, the nineteenth-century authors are far more like each other in writing style than they are like the writers of the Book of Mormon.

Next we present in figure 9 a comparison of the probability results of applying NSC and ENSC to the Book of Mormon texts. It is important to remember that the mathematics in these analyses is not asking, “Who wrote these texts?” The mathematics is asking, “Which texts have the most similar patterns?” The naively applied NSC method estimates that 61 percent of the chapters in the Book of Mormon are most similar in style to texts written by Spalding or Rigdon, without Joseph Smith included as a candidate author. When Joseph Smith is included, the Spalding-Rigdon chapter-attribution proportion drops to

<table>
<thead>
<tr>
<th>Author with the Closest Writing Style</th>
<th>Proportion of Book of Mormon Chapters Assigned to Each Candidate Author</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Using NSC</td>
</tr>
<tr>
<td>Solomon Spalding</td>
<td>22%</td>
</tr>
<tr>
<td>Sidney Rigdon</td>
<td>39%</td>
</tr>
<tr>
<td>Oliver Cowdery</td>
<td>8%</td>
</tr>
<tr>
<td>Parley Pratt</td>
<td>4%</td>
</tr>
<tr>
<td>Isaiah &amp; Malachi</td>
<td>26%</td>
</tr>
<tr>
<td>Henry Longfellow</td>
<td>1%</td>
</tr>
<tr>
<td>Joel Barlow</td>
<td>0%</td>
</tr>
<tr>
<td>Joseph Smith</td>
<td>Not Included</td>
</tr>
<tr>
<td>Someone Else</td>
<td>Not Included</td>
</tr>
</tbody>
</table>

Fig. 9. Comparison of the proportion of Book of Mormon chapters assigned to the candidate authors by each analytical technique. Although the closed-set NSC technique assigns a majority of chapters to Spalding and Rigdon within a constrained set of candidate authors, when allowing for the possibility that the candidate set is incomplete, the open-set ENSC technique assigns an even larger majority of the chapters to an author who was not included in the candidate set—“Someone Else.”
55 percent. But properly addressing the analysis as an open-set problem using the ENSC method, the Spalding-Rigdon proportion of the chapters is a mere 8 percent, and we find that 73 percent of the chapters are attributed to “Someone Else” other than the candidate set of authors.

However, for completeness we need to take the analysis one step further. The control authors are only useful to demonstrate the reliability of the analytical technique. So, after doing preliminary tests, Criddle and associates should have removed the control authors (Longfellow, Barlow, and Isaiah/Malachi) to make their final probability computations. To complete the study properly, after showing the reliability of ENSC, we excluded the control authors, included Joseph Smith as a candidate author, and used only the “early” Sidney Rigdon texts as Criddle and associates should have done. When we did so, we found that NSC assigns only 12 percent to Spalding-Rigdon while assigning 61 percent of the chapters of the Book of Mormon to Cowdery. These results are shown in figure 10. Therefore, even when naively using NSC, Criddle and associates should have concluded that the evidence does not support the Spalding-Rigdon theory without including Cowdery as the primary actor in their theory.

As also shown in figure 10, ENSC assigns 0 percent of the Book of Mormon chapters to Spalding-Rigdon, 4 percent to Cowdery, 3 percent to Smith, and 93 percent to “Someone Else,” indicating that the writing styles of the candidate authors show very little resemblance to the writing styles in the Book of Mormon. Finding that the writing styles of Joseph Smith and Oliver Cowdery are perhaps slightly evident in the Book of Mormon texts is not inconsistent with the claim that Joseph translated the entire book with Oliver acting as his scribe, and that Oliver’s hand transcribed the final manuscript for the printer in preparation for its first publication.

These results confirm that the Criddle and associates study was fatally flawed in concept and execution. Contrary to their contention, stylometric evidence does not provide credible support for the claim that the writing styles exhibited in the Book of Mormon match their candidate authors—Spalding, Rigdon, Pratt, or Cowdery. In fact, the evidence supports the claim that someone other than their candidate authors wrote the book. This is

<table>
<thead>
<tr>
<th>Author with the Closest Writing Style</th>
<th>Proportion of Book of Mormon Chapters Assigned to Each Candidate Author</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Using NSC</td>
</tr>
<tr>
<td>Solomon Spalding</td>
<td>0%</td>
</tr>
<tr>
<td>Sidney Rigdon</td>
<td>12%</td>
</tr>
<tr>
<td>Oliver Cowdery</td>
<td>61%</td>
</tr>
<tr>
<td>Parley Pratt</td>
<td>3%</td>
</tr>
<tr>
<td>Joseph Smith</td>
<td>24%</td>
</tr>
<tr>
<td>Someone Else</td>
<td>Not Included</td>
</tr>
</tbody>
</table>

Fig. 10. Comparison of the proportion of Book of Mormon chapters assigned to the candidate authors with the control authors removed. Without the control authors in the candidate author set, the closed-set NSC technique assigns only 12 percent of the chapters to Spalding and Rigdon, while the open-set ENSC technique assigns nearly all of the chapters to “Someone Else.”
true even when Joseph Smith is considered as a candidate author.

Therefore, based on these findings, we conclude that Criddle and associates have greatly exaggerated their claim to have calculated astronomical odds in favor of Spalding-Rigdon authorship of the Book of Mormon. The results of a properly conducted stylometric analysis are consistent with the Larsen et al. and Hilton results: stylometric evidence does not support the Spalding-Rigdon theory of Book of Mormon authorship.

Although conceptually attractive, NSC classification has limited applicability in stylometric analysis. The ENSC method is far better suited for the analytical challenges faced by researchers investigating open-set attribution questions.

Conclusion

In sum, an authorship attribution study requires the consistent, coherent, and congruent conjunction of historical, biographical, and stylometric evidence to support the conjecture of a writer as the author of a text with disputed authorship. Such a combination of mutually supporting evidence was not set forth by Criddle and associates. Even as a stylometric analysis the Criddle and associates study is invalid since they made a fundamental error in their study design by considering Book of Mormon authorship to be a closed-set problem and then making the logical error of saying the results exclude any other possible authorship, when in fact the researchers had not even allowed for the possibility of other authors in their study design. The open-set possibility is sometimes called the “none of the above” possibility, and in authorship attribution studies an open set is more often the case than not.

The Criddle and associates study used the Nearest Shrunken Centroid (NSC) classification method in an attempt to find evidence in support of the Spalding-Rigdon theory. However, their study design was fundamentally flawed. Although NSC is a sound classification technique, the Criddle and associates study was an unsuitable and mistaken use of the technique. The compounding effect of at least eight major errors rendered their results utterly meaningless.

The paper’s statistical methodology was innovative but misapplied because they failed to realize the need to use an open-set procedure and they did not account for the statistical complications of applying a genomics technique to stylometric analysis.

We conclude that Criddle and associates’ research methodology applied to the long-discredited Spalding-Rigdon theory is fatally flawed and does not provide any new insights into Book of Mormon authorship. Sidney Rigdon did not write the Book of Mormon.

Appendix: Exposing a Methodological Lapse

In an unpublished manuscript, Matthew Jockers attempts to justify the research decision not to include Joseph Smith as a candidate author of the Book of Mormon in the study by Jockers, Witten, and Criddle reported in Literary and Linguistic Computing in 2008.

Appendix: Exposing a Methodological Lapse

In an unpublished manuscript, Matthew Jockers attempts to justify the research decision not to include Joseph Smith as a candidate author of the Book of Mormon in the study by Jockers, Witten, and Criddle reported in Literary and Linguistic Computing in 2008.

32. The contextual evidence does not warrant considering Sidney Rigdon as a viable candidate as the author of the Book of Mormon, as Matthew Roper and Paul Fields discuss in the essay that follows.


well seek to justify this decision since this methodological lapse alone is fatal to the credibility of the published paper. As we will see, this effort at after-the-fact justification on the basis that Smith’s “personal writings reveal a great deal of stylistic variation” is nothing but self-serving special pleading.

First of all, “a great deal of stylistic variation” is hardly a basis upon which to exclude an author as a candidate, especially when that person is listed as the sole “author” on the first printed edition of the book. Further, all historical accounts corroborate Smith’s claim that he dictated the book to scribes word by word. No one else can reasonably be suggested as the prime candidate for authorship. All other candidates must be considered as secondary candidates at best.

Jockers states that “the Smith material is too heterogeneous [highly variable] to be considered a genuine sample of Smith’s style.” High variability certainly makes stylometric analysis difficult, but that cannot be used as a reason to exclude the most likely author as a candidate. If the analytical method is not capable of handling high variability, then Jockers should acknowledge his method’s weakness, abandon it, and find a more capable method. This is how statistical analysis ought to proceed. Further, if he is only interested in studying easy problems, then he should acknowledge his preference, abandon the difficult problem, and find an easier problem to study.

In addition, Jockers offers no basis for what constitutes “too heterogeneous.” The reader wants to ask, “Compared to whom? According to what scale of measurement?” Jockers does not say. Nor does he say why high variability in an author’s style indicates that a writing sample is not genuine. Jockers has no grounds upon which to make such a comparative statement without showing that the other candidate authors had a consistent style while Smith did not, and that the other candidates’ texts are reliable indicators of style while Smith’s are not. He fails to do so, and so this claim is nothing but his own impression of the problem, with no mathematical basis behind it.

With the intent of filling in the gap caused by the exclusion of Joseph Smith as a candidate author in the published study, Jockers compiled in his unpublished study a set of twenty-four documents ranging in size from a mere 112 words to 2,300 words in Smith’s handwriting. These he used to characterize Smith’s style using his previous methodology. As the test set he compiled ninety-six documents ranging from 105 to 10,927 words dictated by Smith to twenty-three different scribes. He added to this set 219 texts by the candidate authors in the published paper—Spalding, Rigdon, Cowdery, Pratt, Isaiah/Malachi, Longfellow, and Barlow.

Using cluster analysis to group texts with similar style and to separate texts with differing styles, Jockers found style variations—as one would expect—among the texts dictated to different scribes. However, he does not show that the relative differences in any way distort or mask Joseph Smith’s style within the documents. So his analysis provides no useful information.

Next he applied NSC using the published study’s closed-set candidate authors. He reports that with a set of 106 words he could distinguish among these authors, but he was able to achieve an error rate no lower than 13 percent. This is an extraordinarily high error rate. For The Federalist Papers he reported a 0 percent error rate. The Federalist Papers is a closed-set problem, as are

35. See the following essay by Matthew Roper, which discusses this historical context.
Joseph Smith’s writings, so the error rates could reasonably be expected to be about the same. This unacceptably and anomalously high error rate should have indicated either that he did not have enough words in the stylistic feature set or that he did not have a set of truly useful distinguishing words. Ignoring this obvious weakness, he proceeded to apply his method anyway. The results are predictable but useless.

For reasons that he does not explain, Jockers considered Spalding, who had died in 1816, when Joseph Smith was still a boy, to be a viable author for Smith’s personal writings. In his results, some of the ninety-six Smith documents were attributed to Spalding, who could not have had anything to do with Joseph Smith’s writings since none of the Rigdon-Spalding theorists have yet managed to bring Spalding back from the dead to compose the Prophet’s diaries. Instead of seeing a big red flag telling him that his method was not informative, Jockers asserts that Joseph Smith was so influenced by Spalding that even his letters to his wife and his diary entries were modeled after Spalding. This is clearly a flawed conclusion. In addition, there is no historical evidence to support the claim that Cowdery, more than being just Joseph’s scribe, was instead the author of Joseph’s writings. But Jockers’s theory requires that Cowdery was more likely the author of Joseph Smith’s known writings than Joseph Smith himself. Apparently any data point, no matter how incongruous, can be marshaled to support a version of the Spalding-Rigdon theory, requiring ad hoc fixes.

Jockers states that there is “a curious stylistic affinity between the style of Spalding and the style of the personal writings” of Joseph Smith. There is nothing curious about it at all. The style measurement is not real. If a “Spalding signal” shows up so prominently (as Jockers claims) in texts that Spalding could not possibly have written, then any assertion that his style is contained in other texts of questioned authorship is obviously invalid. Whatever Jockers measured must have been nothing more than noise. His “Spalding signal” is just “Spalding noise”—to which his biases tune his own ear.

Let’s look at more details in his results. His method says that 14 percent of the ninety-six Joseph Smith documents were written by Spalding, Longfellow, Barlow, or Isaiah/Malachi. Going further, 10 percent were written by dead people—Spalding and Isaiah/Malachi. A method that produces such unreliable results is obviously useless.

Of the ninety-six documents in Jockers’s test set, only twelve can be used to compare the possible effect of Joseph Smith’s scribes on the documents attributed to his authorship. These twelve documents involve other individuals acting as Joseph’s scribe: Cowdery (eight instances), Rigdon (three), and Pratt (one) acted as Joseph’s scribes. For these documents, Jockers’s results indicate the following:

1. None of the twelve were attributed to Joseph Smith. So either Joseph Smith did not dictate any of the documents attributed to his authorship (i.e., personal letters and personal diary entries)—an unlikely scenario—or Jockers’s method is worthless.
2. In the case of Rigdon as scribe, all of his written documents were attributed by Jockers to Pratt. So either Rigdon inexplicably wrote in Pratt’s style when he was Joseph’s scribe or Jockers’s method is not informative.
3. In the case of Pratt as scribe, the only document tested by Jockers was attributed to Pratt. So either a single text of only 123
words (a short paragraph) is too small to reflect anything other than the style of the hand holding the pen or Jockers’s method is not capable of identifying the true author.

4. In the case of Cowdery as scribe, two documents were attributed to Cowdery, two to Rigdon, two to Pratt, and two to other candidates. So either Cowdery was such a literary genius that he could write in his own style and mimic with equal ease the style of two of his friends, plus the style of a renowned poet and two Old Testament prophets for no useful purpose, or Jockers’s method produces meaningless results.

Jockers’s conclusions are an attempt to justify his methodological irregularities by claiming either that (a) Joseph Smith somehow did not write any of the documents (even those written in his own hand) or that (b) his writings are inadmissible as evidence of his personal writing style because of “stylistic variation” and thus need not be considered. If Jockers is admitting that his methodology is incapable of dealing with “stylistic variation,” then he is admitting that his method is inadequate for stylistic analysis.

Paul J. Fields (PhD, Pennsylvania State University) is a consultant specializing in research methods and statistical analysis.

G. Bruce Schaalje (PhD, North Carolina State University) is a professor of statistics at Brigham Young University.

Matthew Roper (MA, Brigham Young University) is a research scholar for the Neal A. Maxwell Institute for Religious Scholarship, Brigham Young University.