

Matthew L. Jockers  
Stanford University

**Title:** Testing Authorship in the Personal Writings of Joseph Smith Using NSC Classification

**Abstract:**

In a co-authored paper (Jockers et al. 2008), my colleagues and I employed both delta and Nearest Shrunken Centroid (NSC) classification in an authorship analysis of the *Book of Mormon*. Our results suggested that several men involved in the early formation of the Mormon Church might have been contributors to the *Book of Mormon*. For reasons detailed in that paper, and summarized again here, we excluded Mormon prophet Joseph Smith in our authorship tests. The work presented here reevaluates our decision to exclude Smith and employs both supervised classification and unsupervised clustering in order to explore the stylistic consistency between documents attributed to Smith (but written in the handwriting of one of Smith's 24 different scribes) and documents in Smith's own hand.

**Introduction:**

In prior research (Jockers et al. 2008) my co-authors and I compiled a corpus of source material from five candidate authors who were either directly involved in the formation of the LDS church or alleged to have contributed to the authorship of the *Book of Mormon*. We did not include Mormon prophet Joseph Smith in our analysis because we determined that there was not enough reliable, authenticated writing by Smith to constitute an ample sample for testing. As Mormon scholar Dean Jessee makes clear in the introduction to *Personal*

*Writings of Joseph Smith* (Smith and Jessee 2002), Smith's speeches, letters, and even journal entries were frequently written by scribes or written in tandem with one or more of his collaborators. In another article that appears in the pages of the "Joseph Smith Papers" online archive (Jessee n.d.) Jessee writes, "only a tiny proportion of Joseph Smith's papers were penned by Smith himself." In many of the documents Jessee collected, we see the handwriting of Smith interwoven with the handwriting of his scribes, sometimes side by side in the exact same letter, journal entry, or document.

Mormon history informs us that Smith frequently used scribes and that he dictated his thoughts to them. Indeed the entire *Book of Mormon* is said to be a verbatim transcript of Smith's dictation. With regard to documenting his visions, thoughts, and experiences, Smith's "philosophy," writes Jessee, "was that 'a prophet cannot be his own scribe.'" That said, on some occasions Smith did put pen to paper, sometimes alone and sometimes in tandem with others. Though Jessee has "attributed" the spirit and content of all of these documents to Smith, the manuscripts show clear physical evidence of other hands at work; thus, the question remains as to whether these scribes were "authoring" or merely "transcribing."

For many scholars, Jessee's research and conclusions provide enough evidence to conclude, as we did in our prior work, that the Smith material is too heterogeneous to be considered a genuine sample of Smith's style. Despite Dean's conclusions, however, many individuals from both within and outside the LDS church either contacted me directly or posted thoughts online suggesting

that we should have tested for Smith in our prior work. Though I remain convinced that the personal writing of Joseph Smith are not reliable for use as samples of determined authorship, these documents do provide fertile ground for another sort of closely related stylistic inquiry and allow an opportunity to investigate the question of whether Smith's various scribes may have contributed more than simple transcription. For this new research I analyze and compare a collection of writings composed by (i.e. independently written, non-dictated works) several of Smith's scribes in order to explore whether the independent styles of these scribes may have been imposed on the works they wrote for Smith in their roles as scribes.

The goal of this work is thus to assess the role (if any) that the scribes may have had in shaping the linguistic and stylistic construction of the documents attributed to Smith. For example, if sections attributed to Smith but written in the hand of Sidney Rigdon are classified as being most similar to the independent writings of Rigdon, such a result would suggest that the role Rigdon played in the dictation process was perhaps more than mere scribe. Alternatively, if the material not in the hand of Smith is classified as being most similar to material that is in his hand, then this would be evidence favoring attribution to Smith and Smith alone—assuming, of course, that Smith himself was not taking dictation.

**Methodology:**

For this analysis, I prepared digital copies of all of the personal writings of Smith that Dean Jessee collected in his text. I then excluded those that only exist in published form--that is I excluded documents attributed to Smith that had

no extant handwritten manuscript as evidence of the document's provenance. Unfortunately, many of the remaining documents were quite short (in word length), so I further excluded documents of less than 100 words<sup>1</sup>. Thus reduced, the corpus of Smith material available for training contained 24 documents in Smith's handwriting. These works ranged in length from 112 words to 2300 words with an average length of 527 (13,172 words total). The test corpus contained 96 additional documents attributed to Smith but in the handwriting of one of 23 different scribes. These works ranged in length from 105 words to 10,927 words.

I compiled all of the text samples into an xml file and coded each sample with a "hand" attribute containing the name of the scribe whose handwriting Jessee had identified. Among the 96 documents there were three documents written in the handwriting of Sidney Rigdon, eight in the hand of Oliver Cowdery and one in the hand of Parley Pratt. In prior research, my co-authors and I had collected primary source material for these three authors. These texts were added to the training corpus. I also added text samples from Solomon Spalding, the Old Testament Books of Isaiah and Malachi, and two control authors, Henry Wadsworth Longfellow and Joel Barlow, again, material collected for prior research.<sup>2</sup> The full training corpus thus included the 24 Smith-written texts from Jessee's collection of Smith's personal writings and 219 text samples from our

---

<sup>1</sup> 100 may still be too small to consider, but I include in Appendix B a table showing the document assignments and the document lengths so that readers may evaluate the assignments in the context of the sample sizes.

<sup>2</sup> The full list of source material can be found in Appendix A of (Jockers et al. 2008).

prior work. Together these 243 texts constituted my training data. The remaining 96 scribe-written documents, my "test" data to be classified.

Utilizing my own tokenization scripts, I lowercased and parsed each text sample to generate a resulting word-feature matrix of dimension 338 rows x 22,308 columns (numbers, such as dates, expressed in numerical form, were removed during this process). My decision to use word features alone was based on the growing consensus that analysis of high frequency words (mostly function, or closed class, words) and/or n-grams provide the most consistently reliable results in authorship attribution problems (Burrows 2002; Grieve 2007; Hoover 2003a, 2003b; Martindale and McKenzie 1995; Y. Zhao and Zobel 2005; Yu 2008). Each row in the matrix represented a single document and each column a single word type, or feature. The cells in the matrix contained the raw counts for the feature in the document. These counts were then normalized to relative frequencies based on document lengths.

To visually explore the Smith materials data before performing classification, I created a sub matrix in which I aggregated the text samples based on scribe: the 24 documents in Smith's hand as one text and the 96 scribe-written texts aggregated by scribe. Using the open source statistics application "R,"<sup>3</sup> I performed a preprocessing step that excluded all word features occurring at a rate of less than 0.1%. This resulted in a new matrix of dimension 24x132. I then performed an unsupervised hierarchal clustering of the data,

---

<sup>3</sup> <http://cran.r-project.org>

using a Euclidian distance measure, to create a histogram showing the relationships between the word usage patterns of each scribe (figure 1).

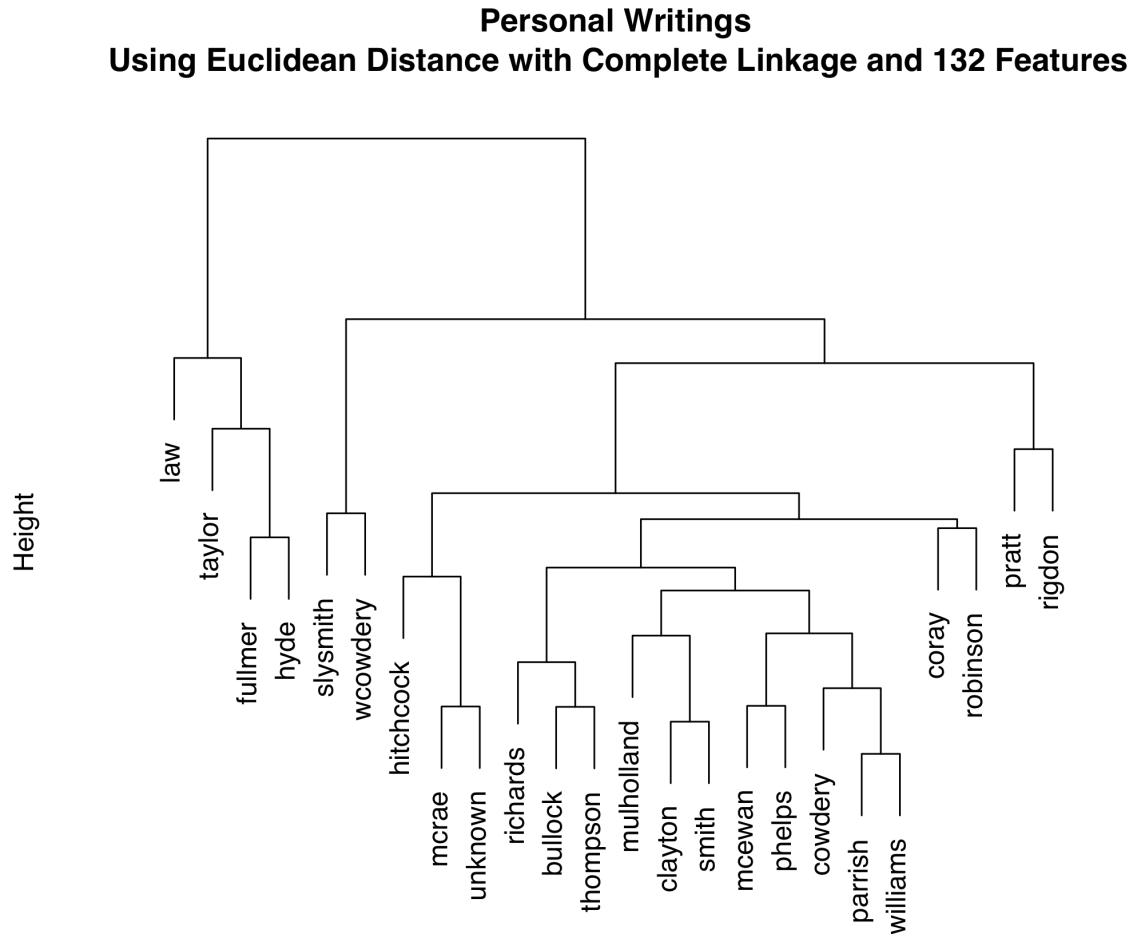


Figure 1.

The documents in Smith's hand clustered most closely with those in the hand of Clayton and Mulholland. The documents written by Pratt and Rigdon showed stylistic affinity.

In order to visualize possible affinities in the word usage patterns of each of the individual texts, I ran a similar procedure without the scribal aggregation. The resulting histogram is too large to reproduce here, but I provide, in figure 2, a snapshot of one branch of the tree. Some clustering by scribe was observed, as was some clustering of the documents written in Smith's hand (e.g. samples labeled "smith" followed by a document type and a primary key referring to the original text sample).

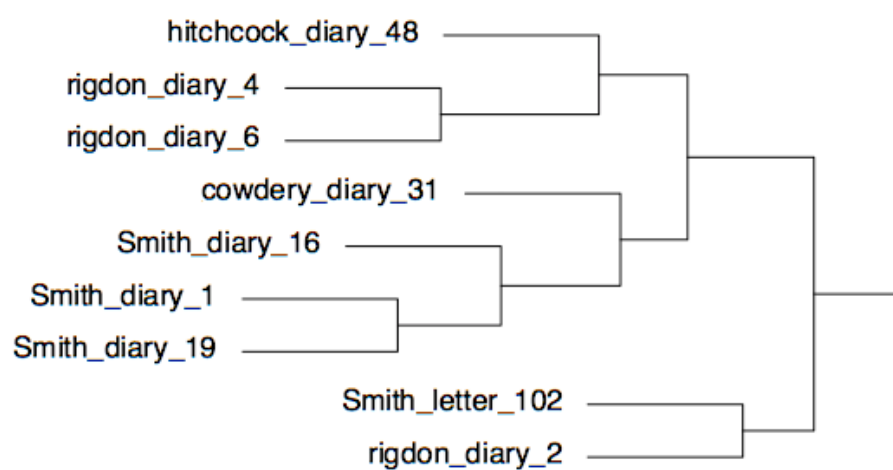


Figure 2

In preparation for the Nearest Shrunken Centroid (Tibshirani et al. 2003) classification of the 96 documents written by scribes, I excluded those words that did not occur at least once in the training sample lexicon from each scribe and also in the lexicon of documents written by Smith. I further excluded words that did not occur with a mean relative frequency, across the samples, of at least 0.1%. Both of these steps were implemented so as to avoid the possibility that some context specific word would have an undue influence on the classification. The resulting list of 106 word-types is found in Appendix A. NSC involves the

use of a tuning parameter to determine the number of features (word-types) to include in the classifier. In order to determine the success rate of NSC at classifying samples of known authorship, and in order to select a value for the tuning parameter, I performed cross validation using a range of values of the tuning parameter. Cross-validation allowed me to estimate the rate of classification error I would obtain if attempting to classify the samples of known provenance. The lowest error rate (12.8%) was obtained using all 106 words. This means that if I used NSC to classify a new sample written by one of the known authors, then the probability of correct classification would be 87.2%.<sup>4</sup>

### **Results:**

In this experiment, NSC misclassified three of the 96 documents to control authors who we know had no connection to Smith or the LDS church. Two documents, one in the hand of Cowdery and one in the hand of Hitchcock, were misclassified as being most similar to Longfellow, a third document in the hand of Richards was misclassified as Barlow. Given the prior speculations about Spalding as a possible contributor to the *Book of Mormon*, I did not specifically treat Spalding as a control text. But I understand that others might wish to think of Spalding as a control, and so I will note that seven of the 96 documents were assigned to Spalding. For those unfamiliar with the *Book of Mormon*, it should be noted that sections of the Old Testament books of Isaiah and Malachi appear either directly or in paraphrased form in the *Book of Mormon*. In our prior work

---

<sup>4</sup> In prior work dealing with the *Book of Mormon*, we observed lower rates of error (8.8% with NSC). Furthermore, in a benchmarking study, Daniela Witten and I recorded 0% error with NSC on the far less complicated Federalist Papers problem (Jockers 2010).



analyzing the authorship of the *Book of Mormon*, we utilized material from the Books of Isaiah and Malachi as candidate material, and that material was thus included in this research as well. In this case, three of the 96 scribe-authored documents were classed as most similar to Isaiah / Malachi.

Depending on one's stance as to whether Spalding is a legitimate candidate author, and depending on how one wishes to treat the Isaiah / Malachi material, these classification results suggest that there were either three or thirteen false positives out of 96 classifications. Of the remaining documents, 32 were assigned to Cowdery, 24 to Pratt, 15 to Smith and 12 to Rigdon.

In addition to assigning a most likely candidate to each text sample, NSC provides a probabilistic ranking of each possible candidate. Given Jessee's contention that these documents attributed to Smith are likely to reflect the style of Smith's scribes and editors, it is useful to examine not simply the first place assignments but the second, third and so on. Table 1 shows a tally of assignments.

	1st	2nd	3rd	4th	5th	6th	7th
Barlow	1	0	5	12	7	25	28
Cowdery	32	21	7	5	7	7	13
IsaiahMalachi	3	1	1	3	7	25	20
Longfellow	2	4	4	10	3	14	26
Pratt	24	12	16	20	16	7	1
Rigdon	12	10	33	12	20	8	1
Smith	15	25	12	12	14	6	7
Spalding	7	23	18	22	22	4	0

Table 1: NSC Rankings

Assuming that Smith worked with only one scribe at a time, it is worth noting in particular the first and second place assignments. In this case, the most prominent signal in the scribe-authored texts is one that matches Oliver

Cowdery. Of the 192 first and second place NSC assignments, 53 (28%) go to Cowdery. Smith is next, with 40 (21%) followed by Pratt with 36 (19%), Spalding with 30 (16%), Rigdon with 22 (11%), Longfellow with six (3%), Isaiah-Malachi with four (2%), and Barlow with one (.05%). Figure 3 provides a graphical representation of the first and second place assignments.

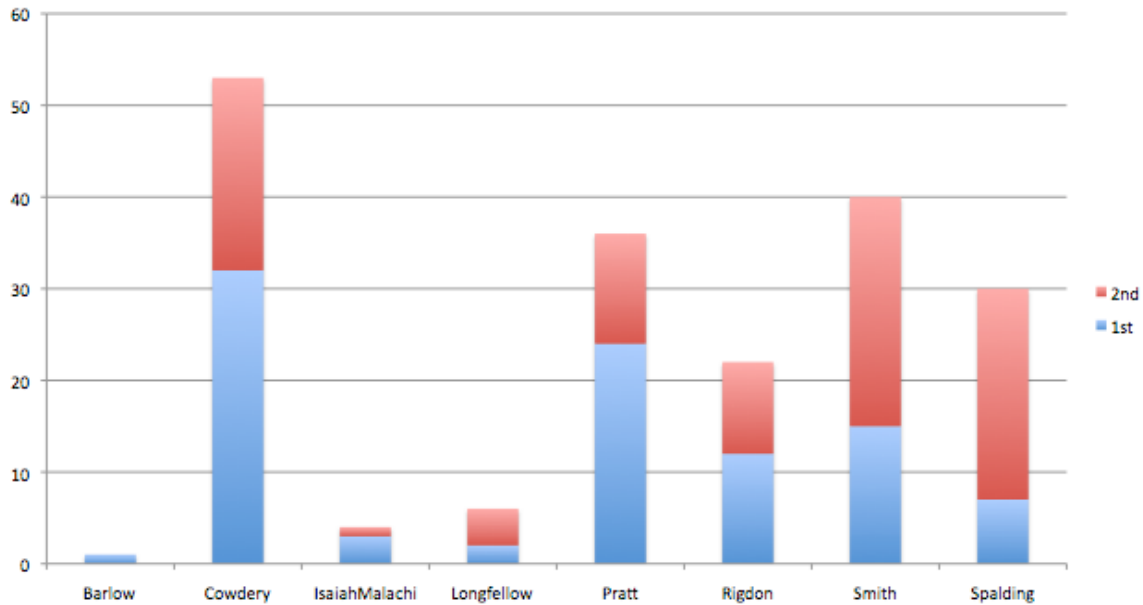


Figure 3. First and Second Place NSC assignments.

Also worth noting are the assignments for the least likely positions of 6<sup>th</sup> and 7<sup>th</sup> place. In this case, the least likely candidates were Barlow with 53 last and second to last place assignments, Isaiah-Malachi with 45, and Longfellow with 40. It makes sense to find these authors represented in the least likely positions since two of these are control texts and the third is a strictly religious document, different from much of the material being explored. Though the personal writings attributed to Smith contain religious content, they are not

specifically religious documents, being instead letters, diary entries and so on.

Figure 4 provides a graphical representation of the least likely candidates.<sup>5</sup>

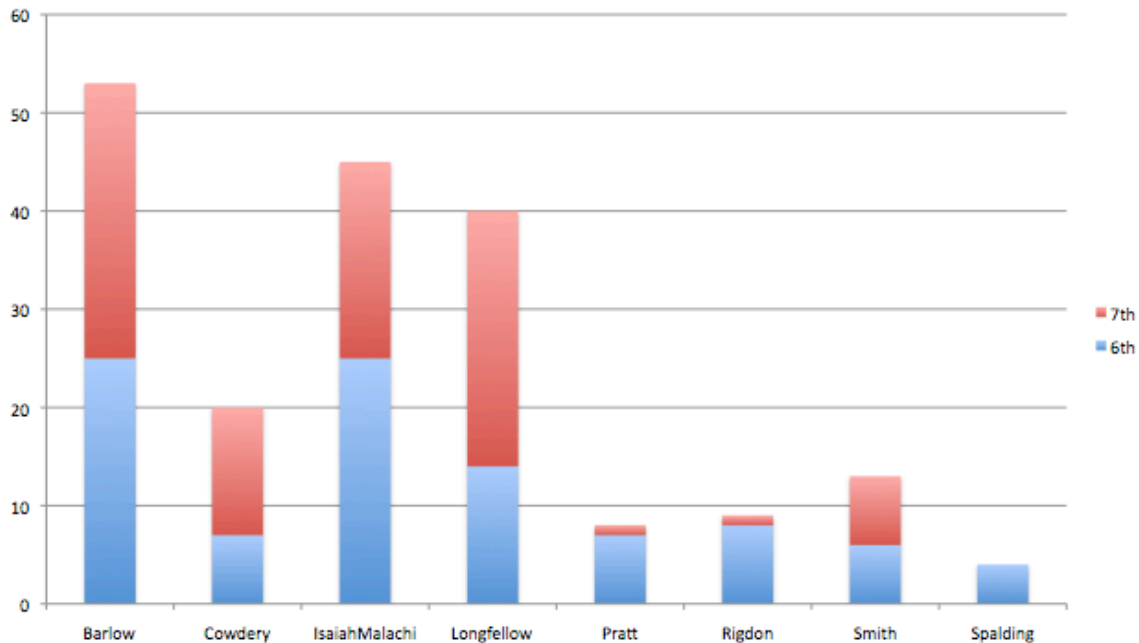


Figure 4: Last place NSC Assignments

## Discussion

The 15 first place (and 25 second place) assignments to Smith appear encouraging. 15 of the 96 personal writings attributed to Smith have a linguistic signal that very closely echoes the signal detected in the documents that are in Smith's hand. Based on this, some might be tempted to imagine a future analysis of the *Book of Mormon* that could test for the presence of Smith by utilizing these 15 documents, along with the 24 documents in Smith's hand, in order to create a justifiable training sample of Smith material. But further thought about these results is warranted. In particular, these 15 documents assigned to Smith represent only 16% of the personal writings not in Smith's hand. In fact,

<sup>5</sup> The taller the column the less likely the candidate.

33% of the scribe-written documents are assigned (first place) to Cowdery and 25% to Pratt. In other words, most of the personal writings attributed to Smith, but not in his hand, are far from stylistically consistent with those in Smith's own hand.

Having noted that, I would add that I find little support for the argument that the scribes (at least those for whom I had non-scribal writings) were imposing their own style upon these documents attributed to Smith. I observe very little correlation between first place assignments to the three candidate authors (Rigdon, Cowdery, and Pratt) who were also Smith scribes and the documents they worked on as scribes. For example, Rigdon is assigned as most likely author for 12 documents, but not one of these was one of the three documents that he penned in the role as scribe. Indeed, the three documents in Rigdon's hand are all assigned to Pratt in first place with Smith in second for one and Spalding for the other two. Of the eight documents in which Cowdery served as scribe, NSC assigns only two to Cowdery in first place and zero in second place. The other 30 first place Cowdery assignments and 21 second place assignments are to documents penned by other scribes. The one document penned by Pratt is assigned to Pratt, but Pratt is also assigned as most likely author for 23 other documents not in his hand.

Of the 15 documents that were assigned to Smith, the scribes most frequently at work are Bullock, Clayton, and Williams. Indeed, three of the five documents penned by Bullock are assigned to Smith, and four of the seven documents penned by Clayton are assigned to Smith and three of the 17 by

Williams. Such a result may suggest that Bullock, Clayton, and Williams were the least intrusive in terms of imposing their own style onto Smith's dictation. To fully test such a hypothesis, however, would require independent writing samples from these authors. A full list of the first and second place NSC assignments is provided in Appendix B.

More interesting than the first and second place assignments alone, however, is the composite picture one gets by looking at the first and second place assignments concomitantly. In the 15 cases where Smith is the first place NSC assignment, he is most frequently paired in second place with Cowdery, (14 times). The only other pairing is with Rigdon, (one time). In the 25 cases where Smith is the second place assignment, he is most frequently paired with Cowdery, in first place, (18 times). Next most frequent is Pratt: when Smith is second, Pratt, appears four times as the first place assignments. Third most frequent is Rigdon: when Smith is second, Rigdon appears twice in first place. If we sum all of the possible first and second place pairings, the most frequently occurring pairs are Smith with Cowdery 32 times, and Pratt and Spalding 20 times. Table 2 provides a full list of the author assignment pairs.

NSC Author Assignment Pairs	Count
Smith and Cowdery	32
Pratt and Spalding	20
Cowdery and Rigdon	11
Cowdery and Pratt	9
Rigdon and Spalding	5
Pratt and Smith	4
Longfellow and Spalding	3
Rigdon and Smith	3
Isaiah-Malachi and Longfellow	2
Pratt and Rigdon	2
Isaiah-Malachi and Rigdon	1
Barlow and Smith	1

Cowdery and Spalding	1
Pratt and Longfellow	1
Smith and Isaiah-Malachi	1

Table 2: NSC first and second place assignment pairings.

The most unexpected result in this analysis may be the strong presence of the Spalding signal. Though Mormon scholars insist that the Spalding theory of *Book of Mormon* authorship is untenable, our prior research detected a strong Spalding signal throughout the *Book of Mormon*. Here again Spalding is prominent, showing up much more frequently in the first (seven times) and second (23 times) place NSC assignments than in the least likely positions of sixth (four times) and seventh (zero time). Spalding could not have directly contributed to the personal writings (he was dead by the time they were written), but the similarity between Spalding's signal and the signal detected in many of the personal writings suggests, at the very least, the presence of a curious stylistic affinity between the style of Spalding and the style of the personal writings attributed to Smith. It warrants noting in this context that Dale Broadhurst has compared the Spalding authored "Oberlin Manuscript" to other works of Mormon history and found some evidence that Spalding fragments may have been used as base text for compositions attributed to Rigdon, including Rigdon's own autobiography as well as some documents attributed to Smith.<sup>6</sup>

### **Conclusions:**

At the end of an experiment such as this, one would hope to have results that are far less messy, results that might lead to an obvious conclusion. The only conclusion I can comfortably reach based on this work is that we were

---

<sup>6</sup> For the Rigdon/Spalding fragments, see (Broadhurst 2009a) and for the Spalding Smith fragments see (Broadhurst 2009b)

entirely justified in not using the personal writings attributed to Smith as training data for our prior work investigating the authorship of the *Book of Mormon*. Smith's collected personal writings reveal a great deal of stylistic variation and the inconsistency certainly seems in harmony with what one would expect from co-authored documents. These results are also consistent with the conclusions of Mormon scholar Dean Jessee, who notes that "the impressions of Joseph Smith given [in the personal writings]. . . probably reflect the personality of the editor[s] more than they do Joseph's" (Smith and Jessee 2002) Based on this research, and on Jessee's, I continue in the belief that the personal writings do not constitute a genuine sample of Smith's linguistic style, and further, that the writings are likely to be exactly what Jessee says they are: documents that reflect the spirit of the man if not his style. Nothing in this current research moves me to change my opinion about the necessity of excluding Smith as an author-candidate in *Book of Mormon* authorship analysis, at least until such time as new, authenticated Smith documents are available. While I do find it tempting to do as suggested above and develop a Smith model based on the texts in his hand and those assigned to him in this experiment, I believe, given the variety of the NSC assignments, that it would be best to resist that temptation. Though it might make for stimulating debate, the leaps of faith demanded would be against this researcher's religion.

## **Appendix A**

106 Words Used in NSC Classification:

a, after, all, am, an, and, any, are, as, at, be, been, before, brother, but, by, came, can, come, could, day, do, down, earth, even, every, for, from, go, god, great, had, hand, have, he, her, him, his, i, if, in, into, is, it, land, let, like, made, make, man, many, may, me, men, more, my, name, no, not, now, o, of, on, one, or, our, out, people, place, said, say, see, shall, should, so, that, the, their, them, then, there, these, they, things, this, thy, time, to, up, upon, us, very, was, we, were, what, when, where, which, who, will, with, would, ye, you, your.



**Appendix B: Ordered by Scribe**

Document	Sample size	First	Second
bullock_letter_127	126	Smith	Cowdery
bullock_letter_128	454	Cowdery	Smith
bullock_letter_134	176	Spalding	Pratt
bullock_letter_135	278	Smith	Cowdery
bullock_letter_91	1011	Smith	Cowdery
clayton_letter_116	682	Smith	Cowdery
clayton_letter_117	456	Smith	Cowdery
clayton_letter_118	305	Smith	Cowdery
clayton_letter_119	645	Cowdery	Pratt
clayton_letter_121	337	Smith	Rigdon
clayton_letter_122	296	Cowdery	Smith
clayton_letter_129	118	Cowdery	Smith
coray_letter_74	269	Rigdon	Cowdery
cowdery_diary_10	853	IsaiahMalachi	Longfellow
cowdery_diary_13	8486	Rigdon	Spalding
cowdery_diary_22	223	Pratt	Spalding
cowdery_diary_25	10927	Rigdon	Pratt
cowdery_diary_26	7649	Longfellow	Spalding
cowdery_diary_31	478	Pratt	Spalding
cowdery_letter_75	661	Cowdery	Rigdon
cowdery_letter_77	2039	Cowdery	Rigdon
fullmer_letter_109	236	Cowdery	Smith
fullmer_letter_111	113	Smith	Cowdery
hitchcock_diary_40	2910	Cowdery	Smith
hitchcock_diary_42	406	Longfellow	Spalding
hitchcock_diary_46	105	Spalding	Pratt
hitchcock_diary_48	2048	Pratt	Spalding
hitchcock_diary_50	169	IsaiahMalachi	Rigdon
hitchcock_diary_52	2237	Pratt	Spalding
hyde_letter_99	2973	Cowdery	Pratt
law_letter_114	601	Cowdery	Smith
mcewan_letter_132	1015	Cowdery	Smith
mcewan_letter_133	135	Spalding	Pratt
mcewan_letter_136	783	Cowdery	Pratt
mcrae_letter_92	153	Rigdon	Cowdery
mulholland_history2_60	520	Pratt	Rigdon
mulholland_letter_79	364	Cowdery	Pratt
parrish_diary_33	5964	Pratt	Smith
parrish_diary_35	341	Cowdery	Pratt
parrish_diary_39	2750	Pratt	Cowdery
parrish_diary_41	669	Pratt	Spalding
parrish_diary_43	780	Spalding	Pratt
parrish_diary_45	1768	Pratt	Longfellow
parrish_diary_47	248	Pratt	Spalding
parrish_diary_49	1005	Pratt	Spalding
parrish_diary_51	1488	Pratt	Spalding
parrish_diary_53	937	Pratt	Smith
phelps_letter_126	1829	Cowdery	Rigdon
phelps_letter_130	2386	Cowdery	Smith
phelps_letter_131	1071	Rigdon	Smith

phelps_letter_83	1154	Smith	Cowdery
pratt_diary_18	123	Pratt	Spalding
richards_letter_112	310	Cowdery	Smith
richards_letter_115	317	Cowdery	Spalding
richards_letter_123	976	Cowdery	Smith
richards_letter_137	3353	Smith	Cowdery
richards_letter_138	778	Cowdery	Smith
richards_letter_139	7559	Barlow	Smith
richards_letter_142	469	Cowdery	Rigdon
richards_letter_143	700	Rigdon	Cowdery
rigdon_diary_2	485	Pratt	Smith
rigdon_diary_4	1378	Pratt	Spalding
rigdon_diary_6	2861	Pratt	Spalding
robinson_letter_84	691	Smith	Cowdery
robinson_letter_85	173	Pratt	Spalding
robinson_letter_86	508	Pratt	Cowdery
sylsmith_diary_44	1261	Pratt	Spalding
taylor_letter_110	192	Cowdery	Rigdon
taylor_letter_140	301	Smith	Cowdery
thompson_letter_100	892	Rigdon	Spalding
thompson_letter_103	367	Cowdery	Pratt
thompson_letter_104	114	Cowdery	Rigdon
thompson_letter_105	504	Cowdery	Smith
thompson_letter_106	1202	Cowdery	Smith
thompson_letter_107	410	Cowdery	Pratt
thompson_letter_108	2844	Cowdery	Smith
unidentified_diary_27	758	Cowdery	Smith
unidentified_letter_90	2261	Rigdon	Spalding
wcowdery_diary_54	660	Spalding	Rigdon
williams_diary_11	740	IsaiahMalachi	Longfellow
williams_diary_21	550	Pratt	Spalding
williams_diary_32	403	Spalding	IsaiahMalachi
williams_diary_34	499	Rigdon	Cowdery
williams_diary_37	121	Pratt	Spalding
williams_history_55	170	Spalding	Longfellow
williams_history_57	145	Pratt	Spalding
williams_history_59	569	Pratt	Smith
williams_letter_61	234	Smith	Cowdery
williams_letter_64	574	Cowdery	Smith
williams_letter_66	844	Rigdon	Smith
williams_letter_68	154	Smith	Cowdery
williams_letter_69	317	Rigdon	Spalding
williams_letter_70	133	Cowdery	Smith
williams_letter_73	206	Smith	Cowdery
williams_letter_76	184	Rigdon	Cowdery
williams_letter_80	229	Cowdery	Smith

## References:

- Broadhurst, D. Access (2009a). *Another Spalding Fragment* 2009a [cited 9/25/2009]. Available from <http://www.solomonspalding.com/bomstudies/fragment.htm>
- . Access (2009b). *Spalding Fragments* 2009b [cited 9/25/2009]. Available from <http://www.solomonspalding.com/bomstudies/fragmnt2.htm>.
- Burrows, J. (2002). 'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 17 (3):267-287.
- Grieve, J. (2007). Quantitative Authorship Attribution: An Evaluation of Techniques. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 22 (3):251-270.
- Hoover, D. L. (2003a). Another Perspective on Vocabulary Richness. *Computers and the Humanities* 37 (2):151-178.
- . (2003b). Multivariate Analysis and the Study of Style Variation. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 18 (4):341-360.
- Jessee, D. Access *Joseph Smith and His Papers: An Editorial View*. Joseph Smith Papers Project [cited 2/25/2009]. Available from <http://josephsmithpapers.org/Essays/Jessee.pdf>
- Jockers, M. L., D. M. Witten, and C. S. Criddle. (2008). Reassessing Authorship in the Book of Mormon Using Nearest Shrunken Centroid Classification. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 23:465-91.
- Jockers, M. L., Witten, D. M. (2010). A Comparative Study of Machine Learning Methods for Authorship Attribution. *Literary and Linguistic Computing*.
- Martindale, C., and D. Mckenzie. (1995). On the Utility of Content Analysis in Author Attribution: The Federalist. *Computers and the Humanities* 29 (4):259-270.
- Smith, J., and D. Jessee. (2002). *The Personal Writings of Joseph Smith*. Salt Lake City: Brigham Young University Press, Deseret Books.
- Tibshirani, R., T. Hastie, B. Narasimhan, and G. Chu. (2003). Class Prediction by Nearest Shrunken Centroids, with Applications to DNA Microarrays. *Statistical Science* 18:104-117.
- Y. Zhao, Y., and J. Zobel. (2005). Effective and Scalable Authorship Attribution Using Function Words. In *Lecture Notes in Computer Science*. Berlin: Springer.
- Yu, B. (2008). An Evaluation of Text Classification Methods for Literary Study. *Literary and Linguistic Computing* 23 (327-343).