Response to Book of Mormon

We thank Professor Croft for his review of our study. We thank *Sunstone* for the opportunity to respond to his comments, to correct some mistakes that occurred in our original article, and to elaborate on some of our thinking.

Many of Professor Croft's points seem to invalidate, or at least blunt, the findings and implications of our study. But we feel that a closer examination of his arguments shows our major conclusions are still valid. In fact, some of his references actually support our results.

Basic Wordprint Assumption

First, Croft quotes some writers who doubt the basic concept of wordprints. We certainly do not claim wordprints to be infallible. We believe that in descending order, fingerprints are most reliable, voiceprints somewhat less reliable, and wordprints even less reliable. Nevertheless, the literature appears to indicate that wordprints do exist in most, if not all, authors. However, the point can be expected to forever remain unproved. Suppose we could show that wordprints do exist for 2,000 contemporary authors. Another researcher might be able to identify a 2,001st author that did not have a distinct, consistent wordprint.

Croft cited studies that indicated some authors do not have consistent word choice. However, these studies were done in a different context than our work. We quote our article (page 227):

> Some previous investigators of authorship identification have oversimplified the problem. Some have chosen a definition of wordprint and then have taken several controversial passages from an author and tested for statistically significant differences in the wordprint between passages. If any statistically significant differences occurred, they assumed different individuals had authored the passages. We believe a larger view must be taken. In addition to comparing several passages written by the same author, we must also compare them with the works of a control group of contemporary authors. Conceivably, an individual author might produce wordprints which differ in a statistically significant manner and yet be consistent within themselves when compared with other authors' wordprints. We have taken this into consideration in our study by including authors who were contemporaries of Joseph Smith.

Therefore, studies indicating that some authors are not consistent in certain types of wordprints do not necessarily invalidate the wordprint concept, unless they also show that these authors cannot be correctly identified when viewed in a larger context.

We give examples to make our point. One author (Yehuda T. Radday, "The Unity of Isaiah: Computerized Tests in Statistical Linguists," unpublished reports, Israel Institute of Technology, 1970, pp. 1-172) has used analysis of variance techniques to claim that the Book of Isaiah was written by two or more people. Radday studied only Isaiah. Later researchers (L. LaMar Adams and Alvin C. Rencher, "A Computer Analysis of the Isaiah Authorship Problem," *BYU Studies* 15 [Autumn 1974]: 95-102; and "The Popular Critical View of the Isaiah Problem in Light of Statistical Style Analysis," *Computer Studies in the Humanities and Verbal Behavior* 4 [1973]: 149-157) considered Isaiah's works in a larger framework. They found that the Book of Isaiah had more internal consistency than any other Old Testament book of the approximate time period and that Isaiah I and Isaiah II (as hypothesized by Radday) were very close in wordprints, as *compared to other* authors of that time period.

Studies indicating that some authors are not consistent in certain types of wordprints do not necessarily invalidate the wordprint concept.

In our Book of Mormon study, we would have preferred to have had other writings in "Reformed Egyptian" written between the period of 600 BC to 400 AD for control texts. Of course we could not find such writings, so we used Joseph Smith and other contemporaries who might have written the book.

Croft claims that we have omitted in one of our sources (A.

"Wordprints" Reexamined

Wayne A. Larsen and Alvin C. Rencher

Q. Morton, *Literary Detection*, New York: Charles Scribner's Sons, 1979) references that do not support our methods. He quotes A. Q. Morton to claim that non-contextual words do not yield stable wordprints. However, Croft fails to note that our study differs from most of Morton's referenced studies. We not only had comparisons within authors, but also comparisons between authors. Strictly speaking, it was not necessary for us to demonstrate a stable wordprint within an individual author, only that the within-author variation was less than the between-author variation. Note also that our non-contextual words worked very well on our control texts.

We only needed to show that the within-author variation was less than the between-author variation.

In another study which considered both within-author and between-author variation (Marvin H. Folsom and Alvin C. Rencher, "Some Characteristics of Subordinate Clauses," *Phonetica* 26 [1972]: 227-234), significant differences were found between authors but the variable examined was relatively consistent from book to book within each author.

Some of Croft's references on this important point actually support our approach. Croft quotes Anthony Kenney as stating, "How far authors are consistent in speech habits such as vocabulary choice is a matter of keen debate." One of Croft's major stylistic sources is the book *Statistics and Style*, edited by Dolezel and Bailey. Werner Winter states in the lead article in this book: "Literary stylistic analysis relies heavily on *features of the vocabulary* and *their distribution*" (p. 5, italics added). We used features of vocabulary and their distribution.

Croft further cites Kenney in an example involving Aristotle, where particles and simple connectives were shown to be inconsistent in Aristotle's writings, sometimes even within a single work. The question we would ask at this point is: If they are not quite consistent within a single work, how do they compare to other authors? They may appear relatively consistent in such a comparison.

Croft cites two papers in the collection by Dolezel and Bailey. A closer examination of one of these articles is very revealing. Croft states: "One selection from this book is by Kai Rander Buch. He shows that the sentence length of an author changes over time." But Croft does not mention that these changes over time were just barely significant statistically. Also the study does not make clear how sentence length might vary from one author to another.

Croft also mentioned two other studies which were specifically discussed by Morton to support Morton's contention that common words are readily influenced by literary form and thus are not good indicators of authorship: studies of the nine books of history by Herodotus and also the Federalist Papers.

According to Croft, Morton showed that the use of the definite article varied among the books by Herodotus. But a careful reading of Morton's presentation of the material shows that he lumped all definite articles together, simply counting all occurrences of any definite article. What such a test means is certainly not clear. Morton concluded that

> the writer's choice of words is not an effective discriminator when it is taken in total as his vocabulary. His choice and rate of using frequent words is much more effective, especially when there are large amounts of text to sample and few contenders for authorship of it. (p. 107)

Frequent words and large amounts of text describe our situation. Morton goes on to advocate the "placing of words" as the best discriminator. We did not use such a characteristic. However, it is evident that Morton does not downplay the use of frequent words to the extent implied by Croft.

We comment further on Croft's quote from Morton on the Federalist Papers study (Mosteller-Wallace). Croft claims that this work is an exceptional situation, but he again ignores his primary source, Bailey. In the closing article in the book *Statistics and Style*, Bailey gives a historical review of statistical stylistics. Interestingly, he spends more time on the Mosteller and

Wallace study than on any other single study and gives it full and unqualified support. He concludes:

> The technique developed in the studies of Ellegard and Mosteller/Wallace can be taken as models for authorship discrimination tests and by extension for other statistical investigations in which characteristic traits of style need to be determined (page 227).

Our wordprint definition is patterned after Mosteller-Wallace, a model for authorship discrimination tests.

In spite of the statement Croft attributes to Bailey, which questions the existence of wordprints, the tone of the book edited by Bailey and Dolezel is supportive rather than critical. We quote from Bailey's concluding statement:

> A variety of cultural and academic trends have combined to inhibit the development of statistical stylistics. . . . Statistical methods concern themselves with the broad tendencies and general characteristics in the data to which they are applied. Moreover, such techniques require a degree of explicitness in assumptions and procedures that is not highly valued in a discipline in which subjective and intuitive judgments are prized. Yet despite these apparent objections, statistical stylistics illustrated both in the works mentioned in this historical sketch and in the essays in this collection deals with questions that are of particular interest to the literary critic who wrote the work. In what directions did this writer develop? What are the constraints imposed on the writer by his language? How does the selection of the mode of presentation influence the shape of his work? Certainly these are all significant questions of vital interest to the literary scholar and statistical methods can help him provide answers to them. The virtues of the techniques arise from their generality and their explicitness (page 231-232).

We are thus led to conclude that these references cited by Croft do not support his thesis at all, but in fact support our own—that statistical stylistics is a reliable tool in the study of authorship. Croft has certainly not made a case from the literature to discredit our methods.

Raw Data Used

Croft believes we should have used the 1830 edition of the Book of Mormon rather than the present edition. All else being equal, we would have preferred using the 1830 edition. However, all else certainly was not equal—the present edition is on computer tape. But is it clear that the earliest edition is always the best? Larry Browning of the BYU Translation Sciences Institute prepared a statement for this rejoinder regarding the problems associated with assuming that the 1830 edition is the most "authentic" version (copy of rejoinder in possession of authors). He pointed out the combination of deliberate changes and accidental errors which have been introduced into the Book of Mormon text. Some spelling mistakes, for example, were introduced by the first scribes who transcribed for Joseph. Other changes crept in as Oliver Cowdery hand copied the manuscript for the printer and as the printer typeset the text. Even some copies of the 1830 version differ, so there is no unique 1830 edition. Some of the textual changes made after 1830 were for the purpose of correcting errors introduced in the 1830 printing. Thus in some aspects the present edition may better reflect the original manuscript.

We must recognize that most of the editorial changes Croft refers to were made in Joseph Smith's lifetime. This editing seems little different from the editing that exists on some of our control texts. It is not clear which edition would be preferable for making internal comparisons *within* the Book of Mormon. Perhaps Croft's comments would be more valid in a comparison of the Book of Mormon to the control texts, but we believe his concern is exaggerated.

Even if one agreed with Croft's concern, his generalizations in this section of his critique (as in other sections) are much too broad. He mentions two words which *might* have significantly different frequencies if we had used the 1830 edition: "that" and "which." He concedes that we reported statistical analyses that do not depend on "which." However, we also reported results that did not depend on "that"—analyses on 42 uncommon words. There were no contradictory results.

We ask, would adding 250 "that's" to the 5715 in our study really make any difference in the final result? The possibility seems remote. One reported analysis was based on only ten words. "That" was one of the ten. We noted in our article that eight of the other nine words, when tested individually, showed significant differences across authors. We could drop "that" with little or no effect on the statistical conclusions.

We note also that in another of our analyses 31 of 38 words were found to have statistically significant differences across authors, and in another analysis 21 of the 42 non-common words were individually significant. This should lay to rest any questions raised about specific words that should have been deleted from our study. In fact, the deletion of 15 or 20 words would have essentially no impact on our study.

In his suggestions for future work, he says "original passages . . . unedited . . . *must* be used." Certainly, using unedited works is desirable. But when Croft uses the word *must*, he eliminates nearly every stylistic study ever done.

Problems of Experimental Design

We note at the outset that Croft's repeated use of the term "experimental design" does not conform to common statistical usage. A better term might be "research design" or "sampling design." The term "experimental design" is usually reserved for experiments where different experimental conditions are assigned to the experimental units obtained previously by sampling.

Croft believes we should have separated the passages he labeled narrative and discourse. We considered this but rejected the idea in this beginning study, because we were already breaking the data down into small sections corresponding to authors. We started by treating each quoted individual as a separate author. We were not sure the data would support this division. Perhaps Mormon, Nephi, Moroni, and the other engravers paraphrased the people they quoted. We did not want to confound testing individual authors as opposed to engravers with testing spoken versus written (discourse versus narrative). However, dividing the data as he suggests would have made a stronger study.

Croft is particularly concerned with the phrase "and it came to pass," which he feels is used primarily in historical or written (narrative) passages. He conjectures that many of our observed differences resulted from this one phrase. We respond to this comment from three directions.

First, is the phrase contextual? Larry Adams of BYU's Institutional Research and Planning prepared the following statement for this rejoinder:

MANOVA itself involves a baseline comparison; there is a serious question whether baseline studies are needed.

It has been conjectured the term "and it came to pass" is a narrative term and therefore contextual in literary writing. This is not the meaning of *contextual* as we use it in style analysis of literary authorship. *Contextual* has reference more to the subject of the discourse than whether or not it is narrative. This is because *contextual* has reference to the subject of the specific sentence itself more than the overall subject of the topic. Since stories are narrative, many if not all of the prophetic discourses tend to be narrative. Prophets often use stories to emphasize or put across a point or doctrine. Ancient prophets used the term "and it came to pass" in discourses. The passage in the following references may be termed discourse: Genesis 9:14 "*And it shall come to pass* when I bring a cloud over the earth, that the bow . . ."; Isaiah 2:2 "*And it shall come to pass* in the latter days, that the mountain of the Lord's house shall be . . ."; Isaiah 4:3 "*And it shall come to pass,* that he that is left in Zion . . ."; Isaiah 24:21 "*and it shall come to pass,* in that day, that the Lord shall punish the host of the high ones . . ."; Isaiah 27:13 "*And it shall come to pass,* that the great trumpet shall be blown, . . ."; Isaiah 65:24 "*And it shall come to pass,* that before they call, I will answer . . ."; Joel 2:28 "*And it shall come to pass* afterward, that I will pour out my spirit upon all flesh . . ." These discourses are somewhat narrative, although perhaps not as clearly so as scriptures that are more historical than discourse, if any could be classified as such.

"Came to pass" can also be translated "it happened that," or "it is that." In a certain sense or meaning, "and it came to pass" has the same position in literary authorship style as "for behold." One author tends to use one of these phrases as an idiomatic expression more than another. For example, the author of the book of Joshua used "came to pass" about three times as often as "behold." The author of the book of Jeremiah, on the other hand, used "behold" about three times as often as "it came to pass." If one were to argue that "behold" is less narrative than "it came to pass," and therefore was used by the author of the book of Joshua less often, that person would then not be able to explain why the author of the book of Jeremiah used "behold" far more often than Isaiah and "came to pass" less than the author of the book of Isaiah when the book of Isaiah in the literary sense is less historical than the book of Jeremiah.

"Behold" could also be classified as a narrative phrase, just the same as "came to pass." Regardless of the context, one ancient prophet would use the phrase more than another in giving the same type of narrative discourse. These terms being idiomatic phrases are used more by some authors of ancient scripture than others for the same subject or narrative discourse.

We say "phrase," but in the original language (i.e., Hebrew) each of these phrases is only a single prefix word. Prefix usage is perhaps one of the most pertinent stylistic elements in determining authorship of Hebrew texts. Prefixes in the Hebrew language constitute a major stylistic element that corresponds to the habit-prone parts of speech in English language. The term "function prefix" describes prefixes that are not pronominal, verbal, and participial. Function prefixes are used in literary authorship style analysis because they tend to identify the habit-prone parts of speech and to be non-contextual. Function prefixes refers to such terms as "and from," and "it came to pass," and "for behold." Regardless of context, these and other phrases are used in literary style analysis of Hebrew texts to identify habit prone parts of speech.

References cited by Croft do not support his thesis at all, but in fact support our own–statistical stylistics is a reliable tool.

Second, Croft's comments on this point caused us to reread one of our references (Burgon, Glade L., "An Analysis of Style Variations in the Book of Mormon," Unpublished M.A. Thesis, BYU, 1958). Burgon's study, though not statistical, is in many ways similar to our own. He considers elements of style across the four major engravers of the Book of Mormon—Nephi, Jacob, Mormon, and Moroni. He splits the text into three divisions. Two of these are similar to Croft's (discourse and narrative); the third is abridgement. For each of his elements of style, he reported at least eight data points: four discourse frequencies and four narrative frequencies. *After making the discourse/narrative distinction, Burgon found wide differences between authors.*

In particular, Burgon investigated "and it came to pass." As Croft conjectures, he found a difference between discourse and narrative, but he also shows a widely differing usage rate for authors within narrative. His data are as follows:

Uses of "it came to pass" Per Thousand Words

	Nephi	Jacob	Mormon	Moroni
Narrative	10.3	7.7	1.2	3.4
Discourse	2.1	0.0	.4	.9
Abridgement	—	—	8.4	12.3

Burgon did not define the writings of these authors exactly as we did, but it appears that this stylistics phrase (even if somewhat contextual) is used differently by different authors. Burgon also investigated the use of "behold" finding a large difference between Moroni and the other engravers within narrative writings.

Third, if "it came to pass" is even possibly contextual, we should have also tested for wordprint patterns without the phrase, and we did. We used 42 non-frequently occurring words (none of the words in the phrase "and it came to pass" are in this list of words). The results were just as indicative of multiple authorship as those that included "and it came to pass."

In summary, Croft's conjecture that many of our results were related to the use of this phrase alone does not hold.

Croft complains about our MANOVA test on the control texts combined with the Book of Mormon texts. His comments here are correct based on the printed version in *BYU Studies*. The MANOVA test that we report in that article is not conclusive, just as Croft suggested. This test was meant to introduce the readers to a set of more meaningful tests. The problem is that some results were inadvertently left out in the printed study. We apologize for this omission in the printed version. The results were in earlier drafts, but somehow were omitted in the final product. We actually performed two of the three tests Professor Croft suggests. We compared the Book of Mormon word frequencies to Joseph Smith's wordprint, and we compared the Book of Mormon word frequencies to Solomon Spalding's wordprint. In both cases we got results of billions to one odds against either of these nineteenth-century figures being the author of the Book of Mormon.

We did not explicitly test Sidney Rigdon versus the Book of Mormon, but we believe our discriminant plots firmly establish his wordprint as completely distinct from the Book of Mormon.

Croft comments on our quote: "It does not seem possible that Joseph Smith or any other writer could have fabricated a work with 24 or more discernable authorship styles." He states correctly that we have demonstrated only at least two distinct styles with the MANOVA test. However, in one of our MANOVA analyses we did compute orthogonal contrasts that demonstrated 15 distinct styles. So although we did overstate the results in that one sentence, results not reported in detail support many different wordprint styles.

Croft's third major statistical comment is in error. He claims that the discriminant analyses results we published are not

what they appear since 65 percent of the book is written by Mormon, and one could obtain a high level of correct classification results just by chance. He is mistaken in the 65 percent figure. Croft quotes the table of engravers, when he should quote the table of writers, where it is clear that only 36 percent of the book was originated by Mormon. So his suggestion that we can obtain 70 percent to 80 percent correct classifications by chance is wrong.

We did not include many detailed statistical results in our article because we were writing to non-statisticians. Since we have been challenged, we now include one more result.

We developed an index to measure how accurately the authors grouped in cluster analysis; that is, if our alleged authors were correctly labeled, how well did the various authors cluster together? We obtained 94.26 for Book of Mormon authors. We then simulated what would happen if our labels were meaningless, that is, if there were only one author for all blocks of words. In this case, any correct cluster would be by chance. We repeated this simulation 100 times. The average index for the simulation was 45.73; the maximum value was 60.32. It is clear that the actual data gave *much* better results than what would have occurred if we did not have distinct wordprints. The classification and cluster results we obtained could not have occurred by chance.

Croft claims that when using authors with only a few blocks of words, it is impossible to test the assumption underlying the multivariate analysis of variance. He insists we need more than four or five blocks of words. However, R. Gnanadesikan in *Methods for Statistical Data Analysis of Multivariate Observations,* (John Wiley & Son, 1977, section 6.3.1), gives a graphical method for testing such assumptions even for a small number of blocks of words.

Croft states that our powerful statistical tools are perhaps so sensitive that they have picked up statistical significance when it does not really exist. This implication cannot be supported by anything in our study and perhaps not from the literature. A very good case in point is the study by Marvin H. Folsom and Alvin C. Rencher. ("Zur Frage Der Sprachlichen Unterschiede in Der BRD und Der DDR," *Deutsche Sprache,* [1977]: 48-55). In this study, 40 characteristics of German literature were compared for various authors in the East and West zones in response to questions about whether there was a divergence in literary style between these two areas corresponding to the political division. MANOVA tests involving the 40 characteristics showed no significant difference. In this case, then, our supposedly sensitive tests did not detect any difference in style between the two political zones. It should not be suggested that MANOVA will go around picking up significant differences any place it is applied. In our study we reported several MANOVA results that did not achieve significance—Isaiah I versus Isaiah II, and Jesus versus Sermon on the Mount.

Croft also suggests that we should have conducted baseline studies. This suggestion is sound; yet it ignores very important considerations. First, MANOVA itself involves a baseline comparison. A highly significant result in MANOVA implies that the differences within an author's works are very much smaller, relatively speaking, than the differences from one author to another author. There is a serious question, then, whether any baseline studies are needed.

The second point which Croft has ignored is one of the strongest of our entire study: *We confirmed our methods on the control texts.* Whatever Croft may wish to say about our methods, when we tried them out on the control authors who were known to be different, they revealed differences among these authors very comparable to the differences we found for the Book of Mormon authors. We therefore argue that our methods were almost completely validated by the results we obtained on the control texts. We do intend at some point in the future to look further at the differences within a given author, such as Mormon. We also note that if our methods were so sensitive, Isaiah should have stood out particularly in the discriminant plots, since Isaiah clearly did not write in the nineteenth century. However, Isaiah did not stand out. Croft ignores this point made in our study.

Conclusion

We do not believe our work, or any work in this area, will be unassailable. This is to ask for more than this science can give. Perhaps editorial pressures and our own enthusiasm caused us to make a few statements that upon closer examination may need some revision. Like any beginning work, it is subject to revisions and re-interpretations as additional data emerge. Yet we believe our study is strong evidence in favor of multiple authorship in the Book of Mormon.

We summarize our rejoinder:

1. Croft claims our wordprint definition is questionable at best. We have quoted the literature, even his own references, to show that our methods are strongly supported.

Perhaps editorial pressures and our own enthusiasm caused us to make a few statements that upon closer examination may need some revision.

2. Croft believes the 1830 edition must be used. We have shown that this may not be the case. We also reason that the differences between the analyses on the 1830 and the present edition would be slight.

3. Croft feels that "and it came to pass" was incorrectly used in our study. We have challenged the generality of his conjecture, and besides, results independent of this phrase support multiple authorship. We agree with his suggested division (narrative and discourse) and will make this change when we extend the study.

4. Croft makes three major statistical criticisms. On two of these, he has a valid point. However, additional results, unreported in our article, resolved the difficulty. On the third statistical criticism, Croft quoted the wrong number, and his argument fails.

5. Croft claims we did not have a baseline study. We reply that the control texts served as a baseline study, and the MANOVA in a sense involves a baseline study.

6. We agree with his suggestions for future improvements.

We do not believe that anyone should base a testimony of the Book of Mormon on our results or anyone else's scientific study. While such studies may help decide the question, "Is the book what it claims?" the final affirmation can only come through spiritual channels well-known to Latter-day Saints. We hope our study does motivate some individuals to study the book itself.

Acknowledgment

We appreciate L. LaMar Adams' and Larry Browning's considerable help on this rejoinder.

WAYNE A. LARSEN and ALVIN C. RENCHER teach in the department of statistics at Brigham Young University.