

Close scrutiny of the methodology of the BYU authorship study reveals several areas which seem vulnerable to criticism

Book of Mormon “Wordprints” Reexamined

D. James Croft

In a recent article appearing in *BYU Studies*,¹ Wayne Larsen, Alvin Rencher, and Tim Layton, specialists in statistics at Brigham Young University, used some highly sophisticated, computerized statistical techniques to examine wordprints of authors in the Book of Mormon. The term “wordprint” was coined by the Larsen group to represent patterns of word usage which authors unconsciously repeat in their writings or speeches—sentence length or frequency of certain common words, for example. The Larsen group maintained that such pat-

and unconscious language patterns. These are words such as *the*, *to*, *which*, *it*, and *for* which are used by everyone, regardless of the context of the writing or the speech. The Larsen group accomplished the exhausting task of assigning every word in the Book of Mormon to its speaker or writer. They then measured the frequency with which each of these authors used the noncontextual words.

Using these frequency counts, they established individual author wordprints. By comparing the wordprints

It is impossible to tell whether differential rates indicate the differences among authors or the differences among editors.

terns can be used to determine which passages of writing belong to one author and which belong to another.

Larsen, Rencher, and Layton focused on the use of certain “noncontextual” words as one measure of unique

of different authors, they tried to answer two important questions. The first was whether the wordprints of authors within the Book of Mormon differ significantly from one another. Such differences, if found, would

support Joseph Smith's claim to have translated the writings of several different ancient writers. The second question was whether the wordprints found in the Book of Mormon match those of any nineteenth century authors, including Joseph Smith, Sidney Rigdon, and Solomon Spaulding, all of whom have been suggested by nonbelievers as possible authors of the Book of Mormon.

After conducting numerous tests and analyses, Larsen and his colleagues reached several important conclusions:

1. None of the Book of Mormon selections resembled the writings of any of the suggested nineteenth-century authors.²
2. Joseph Smith's writing is very distinct from that of the authors of the Book of Mormon.³
3. The implications for translation are that the process was both direct and literal and that each individual author's style was preserved. Possibly, it was given to Joseph Smith word for word.⁴

They concluded their article with this statement: "Our study has shown conclusively that there were many authors who wrote the Book of Mormon."⁵

This optimistic and authoritative conclusion has understandably been greeted with enthusiasm from members of the Church.⁶ The authors delivered a forum address on their findings at BYU, and articles about the wordprint analysis have appeared in various newspapers, in *BYU Today*, and in the *Church News*. The comment of a reviewer in *Dialogue* reveals the confidence placed in the study by many educated Mormons; he points to "a Mormon source whose findings [can] be verified," noting that "impartial computers" have proved odds of 10 billion to 1 against single authorship and odds of 1 billion

The cited conclusions about Book of Mormon selections not resembling nineteenth-century authors are inappropriate.

to 1 against Joseph Smith's authorship of the Book of Mormon.⁷

Before such claims can be accepted as conclusive by skeptics outside the church and scholars within, however, the basically faith-promoting wordprint study must be examined with the same rigor and skepticism reserved for faith-damaging works. Unfortunately, close scrutiny of the study indicates that the encouraging conclusions of its authors may be premature and that several areas of the study seem vulnerable to criticism: basic assumptions made about the notion of wordprints, raw data used in the study, the experimental design used, and presentation of some results. By examining these weaknesses in detail and carefully avoiding them in the future, it may yet be possible to design a study of Book of Mormon author styles which yields conclusions less likely to be challenged.

Basic Assumptions

The Larsen group assumed that wordprints can be used to identify "a piece of writing as belonging to a particular author, just as a fingerprint or voiceprint can be traced to its owner or originator."⁸ Furthermore, according to the study, wordprints can be determined

through statistical analysis of countable features unique to a particular author. This process they called stylometry or stylistics.⁹

To support the validity of these assumptions they described successful examples of using statistical analysis to identify the authorship of works of disputed origin.¹⁰ Citing A. Q. Morton's book *Literary Detection* for support, they stated:

In the literature of stylistic analysis we find many references claiming that for a given author these habits (of style) are not affected by (1) passage of time, (2) change of subject matter, or (3) literary form. They are thus stable within an author's writing, but they have been found to vary from one author to another.¹¹

It appears, however, that some authorities in the field of statistical stylistics have serious reservations about these notions concerning the stability, or even the existence, of measurable style. One of the foremost experts in the field of computer and statistical analysis of style is Richard W. Bailey at the University of Michigan. He has said:

The term "wordprint" is an unfortunate one since it reminds people of fingerprints. We know that fingerprints are valid; voiceprints are somewhat dubious; and we're not sure if "wordprints" even exist.¹²

The reason for Bailey's pessimism is that there are several studies which show an author's style is *not* statistically stable across time, subject matter, and literary form. Larsen and his associates cited none of these studies which question the basic assumptions of their Book of Mormon study. In fact, they omitted the ones reported in their own source on stylistics, A. Q. Morton.

In a recent work, Anthony Kenny was no more optimistic than Bailey: "How far authors are consistent in

speech habits such as vocabulary choice is a matter of keen debate."¹³ He then demonstrated the reason for the "keen debate" by examining Aristotle's use of particles and simple connectives as examples of noncontextual words (he used the term "topic-neutral"). It became clear that Aristotle's use of these noncontextual words was neither consistent throughout his corpus of writings nor even consistent within a single work.¹⁴

Moreover, Aristotle himself said, "The style of written prose is not that of spoken oratory."¹⁵ Modern research has supported this statement. There are numerous studies which have shown there are distinct style differences between written and spoken works by the same author.¹⁶ This finding is particularly relevant for the Larsen study since some Book of Mormon authors are primarily historians (Mormon and Moroni) and use the written form, whereas others are orators (King Benjamin and Samuel the Lamanite) and use the spoken form of communication. Therefore, the possibility exists that some of the statistical differences the Larsen group found might be due to the contrast between written and spoken literary forms.

The major stylistics authority cited in the Larsen article, A. Q. Morton, also pointed out that spoken and writ-

ten styles differ:

It would appear that these common words would make good indicators of authorship if it could be shown that an author used them at a constant rate and individual authors differed in their rates of use. The difficulty in using them as a test of authorship is that their occurrence is too readily influenced by the literary form of the work being studied. . . . The variation in the rate of use seems to be connected with changes in the literary form of the text such as the change from speech to narrative.¹⁷

Morton demonstrated this point by showing that the rate of use of the definite article varied among the nine books of history by Herodotus despite an undisputed single author. Morton emphasized that the cases in which commonly used words provide valid tests of authorship are

A statistical difference in the styles of two passages does not necessarily mean they were written by different authors.

"exceptional situations." He cited one such exceptional case—identifying the authors of the Federalist Papers. The Larsen group used this same case to support their assertions that such identifications are more generally possible.¹⁸

Thus, the very existence of measurable, unique author styles is questioned by people in the field of stylistics. The stability of these styles (if they exist) across time, subject matter, and literary form is a matter of intense debate. Since the Larsen study of the Book of Mormon authors used common words as measures of wordprints and did not allow for style differences between the historians and the orators, its results are subject to the same uncertainty and debate.

Raw Data Used

The fact that Larsen and his associates used edited manuscripts as raw data makes their study vulnerable to still another kind of potentially damaging criticism. It is understandable that the researchers used the current edition of the Book of Mormon since it is stored on computer-readable magnetic tapes which are used in the Translation Services Department of the Church. Even with access to these tapes, the researchers performed a prodigious amount of computer work which involved sorting words from the Book of Mormon and assigning them to authors.

In the process of assigning words to authors, they assumed that when one author described a speech given by another, the first author quoted the second word for word. The validity of this assignment process can be debated. But from a statistical point of view, such an assignment procedure was the correct way to handle the problem of whether speeches were quoted directly or paraphrased. Treating these passages as direct quotes even if they are paraphrased does not introduce false word patterns. But treating them as paraphrases when they are direct quotations *would* introduce nonexistent word patterns.

While using the computer tapes to assign words in this fashion avoided some potentially subtle statistical problems, it introduced others. The major problem with using these tapes as the source for the words (and thus the wordprints) of Book of Mormon authors is that the

current edition of the Book of Mormon is an edited version of the original 1830 edition. The critical question of whether any of the potential nineteenth-century authors actually wrote the Book of Mormon cannot be answered adequately unless *unedited* passages from the Book of Mormon are compared with *unedited* passages from the writings of candidate authors. Nor can multiple wordprints (and thus multiple authors) be established for the Book of Mormon using edited passages. If edited passages are used, it is impossible to tell whether differential rates indicate the differences among authors or the differences among editors.

There are enough editorial changes between the current edition of the Book of Mormon and the 1830 edition to make this a significant consideration. Some parties

have counted close to four thousand editorial changes from the 1830 edition to the present edition of the Book of Mormon.¹⁹ Examination of these changes shows that most of them are minor in nature and involve corrections of grammar or efforts to make the text read more smoothly. The significant point for the Larsen study is the fact that most of these minor changes involve the commonly used, noncontextual words which the researchers used to establish their wordprints.

For instance, the word *that* was listed by the Larsen group as the fourth most commonly used word in the Book of Mormon. It occurs 5717 times²⁰ and was used in most of the tests and results reported in the Larsen study. However, there are over 250 places where *that* occurs in the 1830 edition but not in the present edition. One would have to argue that the rates of deletion for this word are proportional to its rates of usage by each Book of Mormon author if one expects the deletions to leave any underlying wordprints unaffected.

An even bigger problem arises with the word *which*. This word is the eleventh most commonly used word in the present edition. It occurs 1716 times. But in the 1830 edition of the Book of Mormon, *which* was often used in places where *who* or *whom* should have been used. There are over nine hundred changes of this type where *which* no longer appears. That is, the frequency of using *which* is underestimated by more than one-third when we use the present edition of the Book of Mormon rather than the 1830 edition. While this word was not used in all the tests reported in the Larsen study, those in which it was used cannot be considered valid.

The researchers recognized that their use of the current edition was somewhat inadequate. They included an appendix note that "we need to determine what differences are introduced by using the 1830 edition of the Book of Mormon."²¹ However, this uncertainty is not reflected in their unqualified conclusions. More than a passing acknowledgement of this need will be required to satisfy scholars both in and out of the Church. Only tests using the 1830 edition will meet generally-accepted experimental design standards.

The data used to establish the word patterns of nineteenth-century authors had similar difficulties. Some of the passages of these authors were taken from

the *Evening and Morning Star*, the *Messenger and Advocate*, and the *Times and Seasons*, all of which are edited sources. Only part of the passages used to determine Joseph Smith's wordprints were taken from unedited works like his journal or personal letters.

Other problems were associated with the nineteenth-century manuscripts. Richard W. Bailey has pointed out other considerations in addition to those concerning editing:

An author of personal documents (such as a diary or journal) may well adopt a style markedly different from that used in writings for an audience of others. Similarly, the same subject matter may call forth different styles on different occasions, while distinct registers may variously encourage or inhibit the personal mark of style. For these reasons, the documents available to establish the styles of author-candidates must resemble the disputed texts as much as possible, not only in mode but in audience, register, purpose, and time of composition as well.²²

Unfortunately, the Larsen group did not address the possibility of such differences; they compared the historical narrative and sermons of the Book of Mormon

Narrative passages and oratorical passages must be treated separately if the "And it came to pass that" phrase is not to bias the results.

with newspaper articles, journals, and letters of the nineteenth-century authors.

The problems of comparing one set of edited words (from the Book of Mormon) to another set of largely edited words (from dissimilar works of the nineteenth-century authors) with the expectation that statistically subtle (and arguably nonexistent) wordprints remain intact are substantial, at best. Larsen and his colleagues need to do significant work in addressing these issues.

Experimental Design

A number of experimental design problems which seriously weakened the study were left unresolved by the Larsen research group. In some cases, the Larsen group also overstated the possible conclusions pointed to by their statistical test results.

The first experimental design problem concerns the often-used Book of Mormon phrase "And it came to pass that. . . ." The researchers in the Larsen study recognized this phrase created problems for them when they noted in their appendix:

We need to devise better definitions for wordprints using, for example, phrases as well as words. "And it came to pass that" was undoubtedly one word in Reformed Egyptian.²³

However, their comment does not convey an appreciation of the serious biases this phrase created in their study. It has been mentioned that word patterns and styles differ between oratory and narrative works of the same author. A close examination of the words of Book of Mormon authors shows that the historians (like Mormon and, in some places, Nephi, Moroni, and Jacob) used the "and it came to pass that" phrase extensively. The orators (like King Benjamin, Samuel the Lamanite, Alma, Amulek, and, in other places, Nephi, Moroni, and Jacob) did *not*; in some it rarely, if ever, occurs. Thus the inci-

dence of the six words in the phrase "And it came to pass that" is highly dependent on the literary form. An occurrence of any one of those six words in a passage of Book of Mormon material is correlated with an occurrence of the other five words and gives an immediate indication that the author is likely to be a historian. Thus the words in that phrase should not be labeled as "noncontextual." Their appearance as a group gives one an immediate indication of the context in which they are being used and, given that Mormon wrote 65.1 percent of the Book of Mormon, points to the fact that Mormon is the most likely author of the passage in question.

The magnitude of the bias that was introduced by the phrase "And it came to pass that" can be noted by examining the words which the researchers used in the Larsen study. The chart of "Frequently Appearing Noncontextual Words" (see sidebar) shows the 38 words used for most of the tests reported in the Larsen study. These words are arranged in the order of frequency with which they appear in the Book of Mormon. The researchers used only the first ten words from this list in many of their statistical tests; these top ten include four of the

words in the "And it came to pass that" phrase. In other tests involving the entire set of 38 words, the fact that all six words from this phrase are among the 15 most frequently-used words means that some indeterminate portion of the difference in wordprints of authors found by the researchers is solely attributable to this phrase.

A sound research design must overcome this deficiency. Narrative passages and oratorical passages must be treated separately if the "And it came to pass that" phrase is not to bias the results. (Note that the narrative and oratorical passages, not authors, must be treated separately. This is due to the fact that some authors like Nephi, Jacob, and Moroni recorded their words in both literary forms.)

A second group of design problems is related to the statistical methods which were used by the Larsen group and to the way that results from the tests were interpreted. Three statistical methods were used by the researchers: multivariate analysis of variance, often abbreviated MANOVA; discriminant analysis; and cluster analysis. In this paper I will deal with the first two.²⁴

MANOVA tests for homogeneity of groups. The researcher selects several characteristics to be measured for each subject. The characteristics in the Larsen study were the common words and the subjects were the different authors. MANOVA examines the entire set of characteristics to see if the subjects differ from one another. If two or more of the subjects (authors) differ on one or more of the characteristics (words), then the MANOVA test signals a statistically significant result. In other words, the test results suggest the improbability that one subject (author) is responsible for the entire set of characteristics (words). Such a signal alerts the researcher to the need for further analysis of that particular result. He can then use one-way (as opposed to multivariate) analysis of variance which would look at one characteristic (word) at a time. This would show if one or more of the subjects

(authors) differ on the use of one particular word.

The Larsen researchers conducted a number of MANOVA tests for the similarity of wordprints in the Book of Mormon. For one test they used 38 noncontextual words in 341 word blocks (they divided the words of the various authors into blocks containing approximately 1000 words apiece) from 33 authors. Twenty-four of the authors were Book of Mormon authors, including Isaiah, Jesus, and the Lord as quoted by Isaiah. Nine of the authors were nineteenth-century persons. It was reported that "the probability that differences as large as those observed (in the 33 authors' wordprints) could occur by chance is less than 1 in 10 billion." In citing from this, or similar tests, the researchers said that "none of the Book of Mormon selections resembled the writings of any of the suggested nineteenth-century authors."²⁵

Unfortunately, the MANOVA test which was described cannot lead to the conclusion cited. In fact, this particular test, by its very design, could not produce anything but a meaningless result. It was mentioned earlier that the MANOVA test signals a statistically significant difference if it finds that two or more authors differ in the use of one or more words. Since the MANOVA test cited used 33 authors, it was not only comparing the Book of Mormon authors to the nineteenth-century authors, it was comparing the nineteenth-century authors to each other and to Isaiah, Jesus, and the Lord as quoted by Isaiah. It is highly probable that Sidney Rigdon and Joseph, for instance, differ in their use of one or more of the 38 words used in the test. Such a difference would cause the MANOVA test to signal an unusual result. Thus the "1 in 10 billion" difference noted by the Larsen group could have been produced solely by differences between two nineteenth-century authors. Or it might have been produced by differences between the words of Isaiah and those of Solomon Spaulding. Or it might have been produced by differences between any two or more of the 12 authors outside the Book of Mormon. This could all take place with Book of Mormon authors having identical word patterns with one another and with one of the candidate nineteenth-century authors. Thus the cited conclusions about Book of Mormon selections not resembling nineteenth-century authors are inappropriate. Such conclusions could have been reached only with paired comparisons of authors. The fact that the MANOVA statistical test produced results at a highly significant level is both expected and devoid of meaning.

The researchers also reported their results of applying

Even the difference in wordprints for at least two authors cannot be claimed without further tests using more reliable data.

MANOVA to both 10-word and 38-word wordprints for the 21 Book of Mormon authors (this time correctly excluding Isaiah, Jesus, and the Lord quoted by Isaiah through other tests were apparently conducted which included these three as well; specific results of the 24-author MANOVA were not reported). They stated that the differences they observed were "among most of the 21 authors"²⁶ and later commented, referring to their Book of Mormon MANOVA tests, that "it does not seem possible that Joseph Smith or any other writer could have fabricated a work with 24 or more discernible authorship

styles (wordprints)."²⁷

The problem with such reporting is that it implies that all 21 authors differed from each other or that they proved the existence of 21 or more authorship styles. MANOVA and the associated technique used to look at the differences in the use of just one word at a time, one-way analysis of variance, do not have the capacity to show 21 or more differences. The MANOVA tests only showed that several words were used at different rates by at least two of the authors. Furthermore, if a word tests as being used at different frequency rates by one-way analysis of variance, this only means that the author with the highest usage rate differed from the author with the lowest usage rate. It would be possible, then, for 20 of the 21 authors to use a word at exactly the same rate. Given the problems of the "And it came to pass that" phrase and the possible changes in word patterns between literary forms used in the Book of Mormon, even the difference in wordprints for at least two authors cannot be claimed without further tests using more reliable data.

Another test was used which seemed to lend support to the claims of the Larsen group: discriminant analysis. This is a two-stage statistical process. In the first stage, mathematical equations are developed to classify material, in this case to assign passages of text to one of the possible authors. The weights and algebraic signs of each equation were determined by analyzing frequency counts to establish the average profile of each author. In the second stage, the equations based on such profiles were used to assign each block of words considered to the most probable author.

The discriminant analysis results look very impressive at first glance. For instance, the researcher built discriminant equations to classify passages taken from the 21 Book of Mormon authors. When they applied their equations to classify the passages from these authors, they found that 93 percent of the passages were correctly assigned to the authors who wrote them. Since in a way the process is self-verifying, they felt that they had vindicated their method of assigning the authors of each passage in the Book of Mormon, and also vindicated the literalness of Joseph's translation:

Alma's writing is different from Mormon's. Since *all* of Alma's words are taken from Mormon's writings, we can conclude that Mormon copied directly from Alma's writings and Joseph Smith translated literally from Mormon's writings. (p. 241-242) (Italics theirs.)²⁸

The 93 percent success rate, however, was achieved

by reclassifying the passages which were used to build the equations in the first place. A true test of classification accuracy for discriminant equations is made when passages *not used to build the equations* are classified. The researchers recognized this bias and reported that when they applied this latter test, their classification success rate was "consistently in the 70 to 80 percent range."²⁹ Unfortunately, they did not report this unbiased success rate in all their discriminant analysis results.

One might argue that even a 70 to 80 percent discriminant analysis classification rate is proof that distinct

authorship styles exist in the Book of Mormon. But this is not necessarily the case. The accuracy of a classification for authors' passages of text cannot be evaluated without knowing the sample sizes involved.

For instance, assume we have passages of text belonging to four different authors. We desire to assign these passages to the author who wrote them. A classification method which randomly assigns passages to authors would achieve a 25 percent level of accuracy (one correct assignment in four attempts). Another method which gives a higher level of accuracy would appear to be a good assignment method.

But this is not the case if the number of passages used to represent the writing of each author is not equal. If the first author's passages represent 65 percent of the passages used, then the naive method of assigning all passages to that author could achieve a 65 percent level of accuracy, much better than the 25 percent we first assumed. Thus, it is not possible to judge how good the reported 70 to 80 percent classification level in the Larsen study is without knowing the number of sample passages used for each author.

It is difficult to determine in some places what sample sizes the Larsen group used. Tests which involved 21 or more Book of Mormon authors could be criticized on the basis of the smallness of the samples representing authors such as Zeniff, Mosiah, Enos, and Father. In many tests these authors could be represented by only one block of one thousand words. It is somewhat unusual to represent a statistical phenomenon with a sample size of one. This is because the statistical tests used by the Larsen group are based on the differences in wordprints between authors compared to the differences in wordprints within the sampled writings of a single author. For those authors with a sample of only one or two passages, it is not possible to obtain a reading on their internal wordprint variation. In such cases, we must rely on the assumption that all authors have the same rate of internal wordprint variation. However, it is not possible to test this assumption with samples of only one passage. Adequate tests of the assumption cannot be performed unless the number of passages used for an author is substantially more than the four or five passages available for most Book of Mormon authors.

Finally, a word should be said about the powerful statistical procedures used in the Larsen study. Most past work in statistical analysis of style has been done with much less sophisticated statistical tools. Little work has

The accuracy of a classification for authors' passages of text cannot be evaluated without knowing the sample sizes involved.

been done with MANOVA and discriminant analysis. Thus we do not know very much about how these tools react when applied to word patterns of the same author and to word patterns of different authors. It may be that due to their sensitivity they can "find" statistically significant differences in the styles of a single author. Even the simple statistical techniques have found these kinds of differences in some works. Thus a statistical *difference* in the styles of two passages does not necessarily mean they were written by *different authors*.

Baseline studies should be conducted with these

powerful statistical tools to see if they pick up such spurious differences. For instance, several passages of Joseph Smith's unedited writings should be tested to see if MANOVA "finds" statistical differences in them. If it does, then the types or periods of writings that produce these differences must be treated separately in future testing.

These considerable problems in experimental design and in the way results were expressed raise questions about the validity of the conclusions drawn from the Larsen study. We do not know, of course, whether tightening the assumptions made about the notion of wordprints, using unedited materials, strengthening experimental design, and more carefully and cautiously drawing conclusions will alter the results. It may not. But as long as even one type of weakness remains, meaningful conclusions cannot be drawn.

A New Approach to the Same Question

Whatever the weaknesses of the Larsen study, however, it has called our attention to an area of Book of Mormon research that should be examined closely. The next step is to design another stylistics study which avoids the problems which plagued the Larsen study.

First, any study of Book of Mormon authorship styles should deal with unedited materials. The 1830 edition of the Book of Mormon must be used in all sampling for Book of Mormon authors. (Actually, it would be better if the original manuscript were used since Oliver Cowdery made some editorial changes in the copy that went to the printer. Unfortunately, only portions of that manuscript still exist.³⁰) Original passages of writing, unedited for grammar or awkwardness, must be used to represent the work of any nineteenth-century authors to which Book of Mormon passages are compared.

Second, there are several methods of defining wordprints. The frequency with which common words are used by authors has been shown to be unstable as a wordprint definition. Alternative definitions must be examined, especially those which have been shown to be stable for a single author over time, subject matter, and literary form.

Third, it would then be best to divide the tests for differences in authorship styles of Book of Mormon authors into tests which compare narrative passages of one author to narrative passages of another author and tests which compare oratorical passages to oratorical passages.

Fourth, baseline studies should be undertaken using

the wordprint definition selected. That is, the works of Nephi should be examined to determine how stable his writings are, using the selected definition and measuring statistical variation with the powerful statistical tools employed in the Larsen study. If the baseline studies show the need, it may be necessary to use other categories for passages beyond just the narrative and oratorical.

Fifth, in tests which compare Book of Mormon authors to nineteenth-century authors, provisions must also be made to compare narrative passages from Book of

Mormon authors to narrative of the nineteenth-century authors. Like comparisons should be made for oratorical passages. Care must be exercised to insure that comparable types of material are used in all tests.³¹ These tests must also control for biases introduced by "And it came to pass that." In addition, archaic words reflecting the King James language style should not be included in these tests.

Finally, the tests which compare the writings of nineteenth-century authors to Book of Mormon passages must be done in a pairwise fashion. That is, Joseph Smith's writings must be compared to those of the Book of Mormon in one test. A separate test must be used when comparing Sidney Rigdon's writings to the Book of Mormon, and so on.

These considerable problems in experimental design raise questions about the validity of the conclusions drawn from the Larsen study.

Conclusion

Results based on the next generation of wordprint studies may yet provide the encouraging support sought by faithful Church members. Certainly any research done in the future will be indebted to Larsen, Rencher, and Layton, who called our attention to an interesting and challenging area of Book of Mormon study. At the present time, however, given the tentative nature of "wordprints" and given the data and experimental design problems inherent in the Larsen study, it would be best to reserve judgment concerning whether or not it is

Notes

- Wayne A. Larsen, Alvin C. Rencher, and Tim Layton, "Who Wrote the Book of Mormon? An Analysis of Wordprints," *BYU Studies* 20 (Spring, 1980) pp. 225-251.
- Ibid.*, p. 234.
- Ibid.*, p. 242.
- Ibid.*, p. 244. This would markedly alter the concept of translation we typically consider. Word patterns measured by usage rates of common words would certainly not survive through the process of translation in its normal meaning.
- Ibid.*, p. 245.
- The importance of these conclusions to people outside the Church is obvious. However, there is also importance to people inside the Church who are engaged in translating the Book of Mormon into foreign languages. If the words of the Book of Mormon authors are so closely translated from their language into English as to preserve their wordprints, then it would appear that a very literal translation should be used in translating from English to other foreign languages. Efforts to make passages flow more smoothly or read better in the foreign language than they do in English would thus be inappropriate.
- Gary Gillum, "Joseph Smith and Thomas Paine?" *Dialogue* 13 ():138.
- Larsen *et al.*, p. 243.
- Ibid.*, p. 226.
- A. Q. Morton, *Literary Detection* (New York: Charles Scribner's Sons, 1979).
- Larsen *et al.*, p. 227.
- Personal communication to the author dated September 19, 1980. Dr. Bailey is the author of numerous articles on statistical stylistics and the editor of two books in the subject area. He also served as an expert witness in analyzing the writings of Patty Hearst, trying to determine whether she or her captors wrote statements attributed to her during her associations with the SLA. His results were largely inconclusive.
- Anthony Kenny, *The Aristotelian Ethics* (Oxford: Clarendon Press, 1978), p. 71.
- Ibid.*, pp. 90-98. Another interesting work relevant to this debate is *Statistics and Style*, edited by Bailey and Dolzel, American Elsevier, New York, 1979. One selection from this book is by Kai Rander Buch. He shows that the sentence length of an author changes over time. Another selection by F. Neiderich Antosch shows that there are statistical style differences between Parts I and II of Goethe's *Faust*. Other articles in this book also challenge the assertions of Larsen *et al.* concerning the reliability of statistical style analysis in identifying unique word patterns for authors.
- Aristotle, *Rhetoric*, Trans., W. Rhys Roberts, in *Aristotle's Rhetoric and Poetics*, Ed. Friedrich Solmsen (New York: The Modern Library, 1954) p. 14130, 2-4.
- Studies showing differences in spoken and written style include: Joseph A. DeVito, "A Linguistic Analysis of Spoken and Written Language," *Central States Speech Journal* 18 (May, 1967) pp. 81-85; Joseph A. DeVito, "Levels of Abstraction in Spoken and Written Language," *The Journal of Communication* 17 (December, 1967) pp. 354-361; James W. Gibson *et al.*, "A Quantitative Examination of Differences and Similarities in Written and Spoken Messages," *Speech Monographs* 33 (December, 1966) pp. 444-451; and Elbert R. Moses, Jr., "A Study of Word Diver-

sification," *Speech Monographs* 16 (November, 1959) pp. 308-312. Several additional studies could be cited.

"Frequently Occurring Noncontextual Words" (Appendix C in Larsen Study)

Word	Number of Occurrences	Word	Number of Occurrences
and	*10,000	with	1,161
the	*10,000	yea	987
of	8,268	should	945
that	5,717	by	882
to	5,160	as	825
unto	2,955	upon	819
in	2,783	but	818
it	2,665	also	795
for	1,990	from	761
be	1,928	there	668
which	1,716	because	651
a	1,673	these	633
not	1,653	therefore	576
came	1,550	when	576
pass	1,486	if	541
behold	1,456	even	528
all	1,401	into	520
this	1,223	would	494
now	1,178	forth	484

*These counts overflowed the printout field on the computer; therefore, exact counts are not available.

17. Morton, *Literary Detection*, p. 102.

18. *Ibid.*, pp. 102-104. The author wrote A. Q. Morton in Scotland and asked him if he had additional examples in which the frequency of commonly used words changed with literary form. In his personal reply, dated October 28, 1980, he stated, "That the rate of occurrence of frequent words changes with literary form is so generally accepted that no one bothers to give any number of examples. But I will append a few figures." Morton's appended figures showed that the frequent words *the, of, and, to, a, in, his, he, was, and I* were used at statistically different rates in different works of Sir Walter Scott. Additionally, other figures showed that William Shakespeare had significantly different usage rates for the words *and, the, to, of, my, I, in, and that* in works of different literary form. Most of these words were used in the Larsen Study (See Sidebar).

19. Gerald and Sandra Tanner, 3,913 *Changes in the Book of Mormon* (Salt Lake City: Modern Microfilms, 1965).

20. Larsen *et al.*, p. 247.

21. *Ibid.*, p. 251.

22. Richard W. Bailey, "Authorship Attribution in a Forensic Setting," *Advances in Computer-Aided Literary and Linguistic Research* (1979) p. 9.

23. Larsen *et al.*, p. 44.

24. Cluster analysis is not a statistical procedure in the same sense as MANOVA and discriminant analysis. It does not have underlying assumptions which the data must satisfy. The Larsen group's use of cluster analysis would be quite interesting and informative were it applied to the data with the experimental design problems noted previously removed.

25. Larsen *et al.*, p. 234.

26. *Ibid.*, p. 233.

27. *Ibid.*, p. 244.

28. *Ibid.*, p. 241-242. (Italics theirs.)

29. *Ibid.*, p. 238.

30. Richard P. Howard, *Restoration Scriptures: A Study of Their Textual Development* (Independence: Department of Religious Education RLDS, 1969) Chapter 2.

31. In Appendix E of their article, the Larsen group shows the results of applying their techniques to determining authorship of the various "Lectures on Faith." These results are much more convincing than those cited in the article as a whole since in the Lectures-on-Faith study the conditions desired by Bailey were more closely met.



D. JAMES CROFT is professor of Management Science in the Graduate School of Business at the University of Utah.