BS"D Draft. 17 Elul 5759 (29 August 99)

The Nations Sample

By Doron Witztum

Part II: The Range of Stories

Introduction:

BMS [1] assert that they investigated 136 different prefixes in Genesis (including the four prefixes investigated by us in our original experiment. Their list can be found in the Appendix, sec. 1). They report: "Amazingly, three of the four forms used by WRR were the best three in our list of 136!" They write: "This is a rather improbable result; we believe it suggests that Witztum, Rips and Rosenberg cheated and tried many possible prefixes before deciding which four to use." They claim that we actually investigated a large number of possible prefixes and only afterwards concocted a story to justify our use of some of the most successful ones.

1. According to BMS's model [2], for any given text and for any list of prefixes it would be possible to fabricate a plausible explanation of why the most successful ones were chosen. As proof of this they cite the fact that they were able to contrive an explanation for the four most successful prefixes in *WGP* (i.e., four of the 136 mentioned above), and to achieve a result of $5/10^8$ in the permutation test! This experiment shows, so they claim, that contriving rationalizations is a simple matter because there are so many potential stories to choose from.

In this part of the article we will examine their assertions one at a time:

1. Are all explanatory stories really equal?
2. Does the success of our prefixes indeed indicate deception?
3. Is the significance they claim to have achieved accurate?
4. Did their principal story really succeed?

In Chapters I, II, III, and IV we will see that the answer to all four questions is "NO," and we will also see why, in our opinion, their assertions are without any foundation.

2. It also turns out that in the course of compiling their lists of prefixes BMS made a great contribution towards proving that the success of our prefixes does indeed stem from a unique phenomenon in the book of Genesis. An analysis of their list of prefixes shows that there is a significant difference between the success of the prefixes in Genesis versus that of their control text, *WGP*. In Chapter V we will examine the success of the "best quartet" in Genesis versus that of *WGP*. We will see that the success of the "best quartet" in *WGP* is just what one would expect to receive by chance, while the best quartet in Genesis is a case of genuine success. The probability of receiving a similar result by chance would be **0.00042**!

I. The Hierarchy of Stories:

BMS claim that for any four words one could contrive a likely story, and that all explanatory stories are of equal status. This is sheer nonsense. The definition of nationhood (which our explanation revolved around) has been a serious topic of discussion for thousands of years in many different cultures. In the Introduction to Part I of this article I presented excerpts from encyclopedias, dictionaries and traditional sources dealing with this subject. This was not a "contrived" story; but a fundamental discussion of the defining characteristics of nationhood. All other "stories" about aspects of nationality are about issues that are *peripheral and secondary to the essential concept.*

In BMS's list of prefixes one can find terms such as "dance" and "way," etc. Whoever tries to claim that "dance" and "way" define nationhood, or form the basis of a story of the same rank, is assuming that the reader is lacking in sense.

II. The Great "Proof" of Deception:

BMS claim to have found a tremendous proof that WRR cheated. What is their "proof"? They claim that three out of four of the prefixes we selected were the most successful of the 136 prefixes they examined. Their assertion is erroneous on two accounts:

1. The data that BMS present to the reader is misleading. The measures that are relevant here are the ranks of P_1 and P_2 in the permutation test – that is, r_1 and r_2 respectively. These are the only measures of significance we used in the original experiment on the "Nations Sample." (Concerning the "story" BMS told to justify presenting irrelevant data – see the Appendix, sec. 2). Therefore, if BMS chose to assign rankings according to the statistic P_2, they should have done so using r_2. Out of the 136 prefixes examined, the rankings of our prefixes using r_2 were: 1, 3, 9 and 16 (according to BMS's own calculations [3]). According to this data our prefixes enjoyed quite a moderate success.
2. Note that BMS's "proof" is *based on the assumption that there is no real phenomenon*. Therefore all prefixes should be of equal significance (or lack thereof). This assumption is clearly a logical error: our research hypothesis in fact *predicts* greater success for these four prefixes.

This is the height of absurdity! They label a success predicted by the research hypothesis *as proof of deception*!

III. Is the Significance They Report Accurate?

BMS report a significance of $\mathbf{r_2} = 5/10^8$ for the four prefixes they selected. However, by their own admission this is inaccurate. In their article BMS criticized the use of the original randomization test for the Nations Sample. They claimed that the results were meaningless for several reasons particular to this experiment [4]. These objections do not apply to the RPWL test (*Randomization by Permutations of Words' Letters*), see Part I and the Appendix there. For further details see the Appendix here, sec. 3. For this reason, RPWL is the preferred method in this case.

Using the RPWL test the results for BMS's four prefixes (using 100,000,000 permutations) are:

$$\mathbf{r_1(BMS) = 6.67 \text{ X } 10^{-5}}, \qquad \mathbf{r_2(BMS) = 1.22 \text{ X } 10^{-4}}.$$

Changing the test produced an enormous difference in the results for $\mathbf{r_2}$ in *WGP*. In Genesis, by contrast, the difference in the results using the two measurements is small – less that one order of magnitude (see Part I).

BMS found their four prefixes through acknowledged optimization using the original permutation test. In Chapter V we will see what happens when optimization is conducted using the RPWL test.

IV. Did Their Principle Story Really Succeed?

No! As it turns out, the story they concocted for the four most successful prefixes in *WGP* is actually a *failure* – despite their claims to the contrary. The illusory success of their story is entirely based on the omission of a number of prefixes, which BMS neglected to mention. However, the neglected prefixes appear in the source text which BMS claim to have based themselves upon (see the Appendix, sec. 4).

Therefore, there is no story; there is only an arbitrary selection of prefixes, nothing more. It turns out to be not as simple as BMS imagined to contrive a story that works, even if one permits himself to consider all stories of equal status and to ignore the fundamental issue of hierarchy.

V. Without Any Stories:

1. As we mentioned above, BMS published a list of 136 prefixes. One reader of their article, Yosef Beremez, proposed investigating how the "best quartet" out of this list succeeds in Genesis, versus the "best quartet" in *WGP*, thus obviating the need for any explanation at all.

In order to compile the "best quartet" for any text T, one must rank all the prefixes in BMS's list twice: once according to the value P_1, and again according to the value P_2 (prefix with lower P_i-value gets higher ranking). We are interested in the four prefixes that appear at the head of each of the two lists. To determine which is the best quartet of the two, we compare the P_1-value for the *union* of the "P_1 four top prefixes", with the P_2-value for the *union* of the "P_2 four top prefixes". The set with the lowest value is the "best quartet."

As it happens, the P_1 quartet was the lower one both in Genesis and in *WGP*. The best quartet in Genesis, as determined by this procedure, includes the prefixes: עם, כתב, רכב, and שדות ("nation of," "script of," "chariot of," and "fields of").

The best quartet in *WGP* includes the prefixes: אלהי, דמי, שר, and מלל ("god of," "blood of," "prince of," "speech of").

The significance of the statistic P_1 was determined for each of the two quartets using the RPWL test. The results were as follows:

For the best quartet in Genesis: $\mathbf{r_1(G) = 4.0 \text{ X } 10^{-10}}$

Whereas for the best quartet in *WGP*: $\mathbf{r_1(WGP) = 6.16 \text{ X } 10^{-6}}$

The enormous difference between these two results compels us to evaluate the probability of receiving such a small value for $\mathbf{r_1(G)}$. This can be achieved by means of a simulation in which we compare this result with the result randomly expected from an experiment involving the list of 136 prefixes in Genesis. It turns out that the probability of receiving such a low value of $\mathbf{r_1(G)}$ is **0.00042**.

Details on the measurements referred to in this section can be found in the Appendix, sec. 5.

2. BMS made experiments (using the permutation test) [4], in order to compare the distribution of the 136 prefixes in Genesis, with their distribution in 10 similar texts. It is clear from their data that the results for Genesis are significantly exceptional. Details can be found in the Appendix, sec. 3.1-2.

Appendix to Part II

1. The list of 136 prefixes:

Here is the list of 136 prefixes as published by BMS in Table 2 of their article. The prefixes are arranged alphabetically:

1. הקידומת הריקה. 2. אב. 3. אבי. 4. אדמת. 5. אוכל. 6. אומת. 7. אופי. 8. אחזת. 9. אי. 10. איי. 11. איל. 12. אילי. 13. איש. 14. אכל. 15. אל. 16. אלה. 17. אלהי. 18. אלוף. 19. אליל. 20. אלף. 21. אלפי. 22. אמת. 23. אנשי. 24. ארץ. 25. אשם. 26. אשמת. 27. אשת. 28. בגד. 29. בגדי. 30. בירת. 31. בית. 32. בן. 33. בנות. 34. בני. 35. בת. 36. בתי. 37. גבול. 38. גבור. 39. גבר. 40. גברי. 41. גדוד. 42. גוי. 43. גזע. 44. גר. 45. גרי. 46. דבור. 47. דגל. 48. דגלי. 49. דין. 50. דיני. 51. דם. 52. דמי. 53. דרך. 54. דרכי. 55. דת. 56. דתי. 57. זהב. 58. זכות. 59. זמן. 60. זמר. 61. זמרת. 62. זרע. 63. חבל. 64. חטא. 65. חטאי. 66. חיל. 67. חילי. 68. חק. 69. חקי. 70. חרב. 71. טבע. 72. כהן. 73. כהני. 74. כוח. 75. כח. 76. כחות. 77. כסף. 78. כספי. 79. כתב. 80. כתיב. 81. לאם. 82. לבוש. 83. לשון. 84. לשן. 85. מאכל. 86. מושב. 87. מזל. 88. מטבע. 89. מכתב. 90. מלאך. 91. מלך. 92. מלכי. 93. מלכת. 94. מלל. 95. ממשל. 96. מנהג. 97. מס. 98. מסי. 99. מעות. 100. מקום. 101. מראה. 102. נמוס. 103. נס. 104. נסי. 105. נשות. 106. נשיא. 107. סוף. 108. סמל. 109. סמלי. 110. ספר. 111. עוון. 112. עון. 113. ע״ז של. 114. עיר. 115. עם. 116. ערי. 117. פשע. 118. פשעי. 119. צאצא. 120. צבא. 121. ראש. 122. ראשי. 123. רוח. 124. רכב. 125. רקוד. 126. שדה. 127. שדות. 128. שיר. 129. שירי. 130. שם. 131. שמות. 132. שפת. 133. שר. 134. שרי. 135. שרש. 136. תושב.

Note – I have reproduced BMS's list here verbatim (including the mistakes).

2. Do BMS know how to contrive a story?

And how! As it turns out, BMS have a special talent for concocting stories. Two of them (BM) were co-authors of an article [5] describing how (almost) all of the choices we made in our original "Famous Rabbis" experiment were biased in our favor. One of the obvious flaws in their analysis was their choice of statistical tests. They chose to check the bias in our choices by calculating a statistic that was only conceived two years *after* the original experiment was performed (i.e., the rank in the permutation test) instead of calculating the measure of success that was actually used at the time of the experiment! When faced with criticism over this [6] they "concocted" a story [7] to "explain" *post hoc* why they did something so peculiar. (The reader may have already guessed that had they used the more relevant statistic, their analysis would have in fact demonstrated the sincerity of our work).

In their article concerning the Nations Sample BMS demonstrate their impressive acrobatic ability. They managed to do precisely the reverse of what

they did in the article mentioned above. The experiment on the Nations Sample really *was* conducted using the permutation test (which had already been applied by then to the second list of Famous Rabbis). Yet this time they chose to base their investigation of bias using the *old* measure, the one that was used in the original work on the Famous Rabbis sample! Does this seem strange to you? Don't worry – they have a story to explain this as well....

They claim [8] that we could not have used the permutation test for our alleged optimization because we did not have sufficient computing facilities. They support this claim by noting that in our original article on the Nations Sample we write that we were unable to perform the billion permutations necessary for calculating the ranking of P_1.

One cannot deny that this story is very cute, but does it reflect reality? Unfortunately it does not. They did indeed quote us accurately, but they left out a minor detail that BMS should be made aware of [9]: The program for calculating the rankings of P_1 was 45 times slower that the one for reckoning the P_2 rankings. It was because of this slowness that we decided to forgo calculating the ranking of P_1 out of a billion permutations. But regarding the much quicker program for establishing the ranking of P_2, the statistic that BMS claims we were unable to calculate, there was no *technical* obstacle to the running of repeated experiments of this type (had it occurred to us to do anything so stupid). Therefore this is the test that BMS should have used were they genuinely looking for evidence of bias.

As I mentioned earlier, when it comes to concocting stories, Scheherazade did not come up to their ankles.

3. Concerning RPWL as the preferred measuring device:

In part I of this article and in the Appendix there, I brought the necessary background for RPWL method and described how it was applied to the Nations Sample. At this point I'm not going to argue with the speculations that BMS raised in order to undermine the original permutation test. Partly because they raised peculiar ideas (for instance, they wrote that there is a dependency between the ELSs of "עם גמר" and the ELSs of "כתב גמר", etc.) and partly because they admit to having no true explanations. The only issue that may be taken seriously (if their data is correct) is their claim of making certain experiments, and getting results that indicate certain anomaly in the distribution of the rankings in the original permutation test.

1. BMS' data concerning the original permutation test:

BMS report in Section 7.1, that they investigated the distribution of the rankings of the 136 prefixes in Genesis. The 136 rank orders were put in 10 bins 0,1,….,9, where bin i contains permutation rank orders $100000i+1$ through $100000(i+1)$ out of a million. They observed "non-uniformity" and "an excessive number of ranks in the smallest bin", and that the minimum "appear to be smaller than expected". To understand the situation better, they created texts equal in length to Genesis, that were supposedly very similar to Genesis. BMS report:

"To investigate whether non-uniformity and exaggerated extremes are the norm, we ran the same tests on 10 control texts. We generated each control text by randomly permuting the order of the words within each verse of Genesis except for the verses in Chapter 10. Chapter 10, where all nation names are found in a specific fixed order, was left alone in order to make the comparison with Genesis and WGP more meaningful." To ensure independence of the results, ELSs lying

entirely inside the section of the text, containing Chapter 10 of Genesis were ignored.

The results for $\mathbf{r}_2$ are tabulated below. Text G is Genesis. The first ten rows show the numbers of prefixes for which the rank order lies in each of bins 0,1,...,9. The last three rows show the smallest, average, and maximum rank orders out of a million.

Bin	G	Text 0	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	Text 7	Text 8	Text 9
0	23	15	19	22	23	18	17	14	13	19	19
1	17	13	13	8	15	12	6	13	18	13	18
2	9	17	13	16	15	8	18	13	15	11	13
3	13	13	14	7	9	7	16	18	10	12	11
4	11	9	7	13	7	10	12	9	13	7	7
5	12	16	11	13	12	16	5	12	8	10	8
6	10	10	7	9	22	22	10	8	13	18	10
7	14	9	13	20	10	9	13	11	15	16	10
8	16	13	15	15	14	14	23	15	13	11	13
9	11	21	24	13	9	20	16	22	18	19	27
min	493	5314	3991	1584	608	874	4979	180	6184	6247	2481
mean	463125	504905	517197	496493	458654	534142	519963	522828	505910	512851	514246
max	990921	999807	999923	999141	997583	999864	999849	999229	997025	991129	998778

BMS conclude:

"The most obvious characteristic of these distributions is their inconsistency. A few are near-uniform, but others are skewed markedly in the positive or negative direction. Both the minima and maxima appear exaggerated.

Several things should be clear from these results. Firstly, WRR's assumption of uniformity in the rank orders is unfounded, as many of the texts give profoundly non-uniform distributions. More importantly, the non-uniformity may be more pronounced at the extremes of the distribution where WRR measure their 'significance levels'."

2. Comment on BMS' Data:

BMS arrived at the conclusion that "the non-uniformity may be more pronounced at the extremes of the distribution", both by checking occupation numbers of bins 0 and 9, and by measuring the minima and maxima.

Looking at the same table and checking the tendency to low ranks at the extremes, we can't escape the conclusion that the data for Genesis are exceptional. It is pronounced in each of the following parameters:

- The occupation number of bin 0 is the highest (there is another such number) of all 11 texts.
- The occupation number of bin 9 is second to the lowest of all 11 texts.
- The minimum is second to the lowest of all 11 texts.
- The maximum is the lowest of all 11 texts.

These exceptional results for Genesis complete the picture we already have from the measurements described in Chapter V above.

3. Data concerning the RPWL test:

In order to make a similar experiment using RPWL method, we created 10 texts following the description of BMS. To save time, (this method needs much more computations), we first checked the extremes. We checked bin 0 and bin 9; (the occupation numbers were found using the extreme 30 prefixes at each edge). We also measured the minima and maxima for each text. The results for $\mathbf{r}_1$ are tabulated below:

Bin	Text 0	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	Text 7	Text 8	Text 9
0	13	14	8	5	13	12	15	14	14	9
9	10	16	20	8	11	13	12	11	16	13
min	2177	16145	3064	34204	14822	2867	5334	2984	7771	13605
max	985212	998915	992263	977197	998763	996302	997474	980781	997363	994129

The results from a similar experiment done for $\mathbf{r}_2$ are tabulated below:

Bin	Text 0	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	Text 7	Text 8	Text 9
0	14	17	14	9	12	15	12	10	10	11
9	12	24	15	7	14	13	13	12	13	15
min	5112	17038	7705	21471	11176	6775	1076	1659	10225	10626
max	984823	996099	998372	999790	999204	997275	995230	989584	995961	998581

We can clearly conclude from these tables, that by using RPWL method we have

- no accumulations at the extremes, and
- no exaggerated minima or maxima

in the control texts.

4. BMS's main story:

BMS's main story, "Celestial Guardians" [10] is based on Nachmanides commentary to Chapter 18 of Leviticus. There, as BMS put it, "he discusses the celestial beings who represent and supervise the Nations of the world".

1. BMS claim that in the course of this discussion Nachmanides uses the descriptive terms: שר (prince), מלכי (kings of), אלהים (gods), and עירין (angels).
2. BMS contrived a set of "guidelines," which they claim leads to the selection of these four prefixes: 1. מלכי, 2. שר, 3. עיר, 4. אלהי.

By examining the relevant passage we can see right away that:

1. Nachmanides actually used a much broader range of descriptive terms, of which BMS arbitrarily selected four.
2. Their supposed "guidelines" are a nothing more than a joke.

Therefore, they have no explanatory story; just an arbitrary choice of prefixes.

1. Nachmanides descriptive terms:

We will present here two excerpts from Nachmanides commentary (Leviticus 18:25), and we will highlight the relevant terms:

"ותטמא הארץ ואפקוד עונה עליה ותקיא הארץ - החמיר הכתוב בעריות, בעבור הארץ שתטמא בהן ותקיא הנפשות העושות, והנה העריות חובת הגוף ואינן תלויות בארץ, אבל סוד הדבר בכתוב (דברים לב, ח-ט) "בהנחל עליון גוים בהפרידו בני אדם יצב גבולות עמים וגו' כי חלק ה' עמו וגו'". והענין כי השם הנכבד ברא הכל, ושם כח התחתונים בעליונים, ונתן על כל עם ועם בארצותם לגוייהם כוכב ומזל ידוע כאשר נודע באצטגנינות . וזהו שנאמר (דברים ד, יט) "אשר חלק ה' אלהיך אותם לכל העמים", כי חלק לכולם מזלות בשמים, וגבוהים עליהם מלאכי עליון נתנם להיותם שרים עליהם ,כענין שכתוב (דניאל י, יג) "ושר מלכות פרס עומד לנגדי ,וכתיב (שם, פסוק כ) "והנה שר יון בא", ונקראים מלכים כדכתיב (שם, פסוק יג) "ואני נותרתי שם אצל מלכי פרס". והנה השם הנכבד הוא אלהי האלהים ואדוני האדונים לכל העולם ,אבל ארץ ישראל אמצעות הישוב היא נחלת ה' מיוחדת לשמו, לא נתן עליה מן המלאכים קצין שוטר ומושל בהנחילו אותה לעמו המיחד שמו זרע אוהביו, וזהו שאמר (שמות יט, ה) "והייתם לי סגולה מכל העמים כי לי כל הארץ", וכתיב (ירמיה יא, ד) "והייתם לי לעם ואנכי אהיה לכם לאלהים", לא שתהיו אתם אל אלהים אחרים כלל. והנה קידש העם היושב בארצו בקדושת העריות וברובי המצות להיותם לשמו, ולכך אמר (להלן כ, כב) "ושמרתם את כל חוקותי ואת כל משפטי ועשיתם אותם ולא תקיא אתכם הארץ", וכתיב (שם, פסוק כד) "ואמר לכם אתם תירשו את אדמתם ואני אתננה לכם לרשת אותה אני ה' אלהיכם אשר הבדלתי אתכם מן העמים" יאמר כי הבדיל אותנו מכל העמים אשר נתן עליהם שרים ואלהים אחרים, בתתו לנו את הארץ שיהיה הוא יתברך לנו לאלהים ונהיה מיוחדים לשמו .והנה הארץ שהיא נחלת השם הנכבד תקיא כל מטמא אותה ולא תסבול עובדי ע"ז ומגלים עריות. והפרשה הזאת הזכירה המולך לכלול עבודה זרה עם זכרון העריות. ועל כולם אמר "אל תטמאו בכל אלה כי בכל אלה נטמאו הגוים ,ותקיא הארץ את יושביה", וכן אמר בפרשה השניה (להלן כ, כו) "ואבדיל אתכם מן העמים להיות לי", שהוא חומר איסור עבודה זרה, ולכך אמר כי מפני שהם מיוחדים לשמו בעבור כן נתן להם הארץ ,שנאמר (שם, פסוק כד) "ואומר לכם אתם תירשו את אדמתם ואני אתננה לכם לרשת אותה אני ה' אלהיכם אשר הבדלתי אתכם מן העמים". והנה בחוצה לארץ, אע"פ שהכל לשם הנכבד, אין טהרה בה שלימה ,בעבור המשרתים המושלים עליה והעמים תועים אחרי שריהם לעבוד גם אותם . ולכך יאמר הכתוב (ישעיה נד, ה) "אלהי כל הארץ יקרא", כי הוא אלהי האלהים המושל על הכל והוא יפקוד בסוף "על צבא המרום במרום" להסיר ממשלת העליונים ולהרוס מערכת המשרתים ,ואחרי כן יפקוד על "מלכי האדמה באדמה". וזהו ענין הכתוב שאמר (דניאל ד, יד) "בגזרת עירין פתגמא ומאמר קדישין שאלתא". יאמר כי הדבר ההוא הנגזר על נבוכדנצר היא גזרת עירין פתגמא ומאמר קדישין שאלתא שגזרו על הכוחות הנאצלין מהן לעשות כך, ויקראו עירין כי מאצילותן יתעוררו הכחות בכל הפעולות, כמו שאמר (שם, שם, י יא) "ואלו עיר וקדיש מן שמיא נחית קרא בחיל וכן אמר גודו אילנא וגו'". ומאמר קדישין שאלתא, כלומר ששאלו מה הרצון העליון עליו ואחרי כן גזרו להעשות כן, וזהו שאמר לו דניאל (שם, כא) "וגזרת עילאה היא", כי הכל מאתו יתברך. והנה השם הנכבד יתברך אלהי האלהים בכל העולם ואלהי ארץ ישראל שהיא נחלת ה', וזהו טעם "וזנה אחרי אלהי נכר הארץ" (דברים לא, טז), כי האלוהות נכרים בארץ השם ובנחלתו, וזהו שנאמר (מ"ב יז, כו) "לא ידעו את משפט אלהי הארץ וישלח בם את האריות והנם ממיתים אותם כאשר אינם יודעים את משפט אלהי הארץ". והנה הכותיים לא היו נענשים בארצם בעבדם את אלהיהם לשלח בהם את האריות, ובבואם בארץ השם ועשו שם כמעשיהם הראשונים שלח בהם האריות הממיתים אותם. וכן שנו בספרא (קדושים יא, יד) "ולא תקיא הארץ אתכם וגו' -- ארץ ישראל אינה כשאר ארצות ,אינה מקיימת עוברי עבירה". ובספרי (האזינו שטו) "ואין עמו אל נכר (דברים לב יב) -- שלא תהא רשות לאחד משרי האומות לבא לשלוט בכם, כענין שנאמר ואני יוצא והנה שר יון וגו'". והוא מאמרם (כתובות קי ב) "כל הדר בחוצה לארץ דומה כמי שאין לו אלוה, שנאמר (להלן כה, לח) לתת לכם את ארץ כנען להיות לכם לאלהים, ואומר (ש"א כו, יט) כי גרשוני היום מהסתפח בנחלת ה' לאמר לך עבוד אלהים אחרים"...

"...והענין הזה הוא במקומות רבים בכתובים, ותראנו מפורש בהם אחר שפקחתי בו עיניך. וכתב ר"א בפרשת וילך (דברים לא, טז) "ידענו כי השם אחד והשינוי יבא

מהמקבלים, והשם לא ישנה מעשיו כי כולם הם בחכמה, ומעבודת השם לשמור כח הקבול כפי המקום על כן כתוב את משפט אלהי הארץ ,ואמר ביעקב הסירו את אלהי הנכר, והפך המקום הדבק בעריות שהם שאר והמשכיל יבין." אלו דבריו ז"ל. ואל תשיב עלי מפסוק "מיכאל שרכם" (דניאל י, כא) כי הוא שר משרת לבקש רחמים על ישראל ,לא שר מלכות וממשלה .וכן היה שר צבא הנראה ליהושע ביריחו (יהושע ה, יג) הראה לו כי השם שלחו ללחום מלחמותיהם כענין בחזקיהו (מ"ב יט, לה), וגם שהיה זה בהיותנו בחוצה לארץ..."

In all, we have highlighted 26 relevant terms. Here they are in the order in which they appeared:

1. כוכב star. 2. מזל constellation. 3. מזלות constellations. 4. מלאכי the angels of . 5. שרים lords. 6. שר מלכות the prince of the kingdom of (a prince exercising royalty). 7. שר the prince of. 8. מלכים kings. 9. מלכי the kings of. 10. אלהים gods. 11. אדונים lords. 12. מלאכים angels. 13. קצין chief. 14. שוטר observer. 15. מושל ruler. 16. משרתים servants. 17. מושלים rulers. 18. עליונים celestial powers. 19. עירין the wakeful ones. 20. קדישין the holy ones. 21. עיר a wakeful one. 22. קדיש a holy one. 23. אלהי the gods of. 24. אלוהות gods. 25. אל the god of . 26. שרי the princes of.
(The translation of the 26 terms was taken from the same source used by BMS.)

2. Contriving guidelines:

BMS explain that they arrived at their four prefixes using three guidelines:

1. *They only selected terms that are used explicitly by Nachmanides.*
   - In fact, as we have seen, Nachmanides actually uses explicitly many more terms than the ones they mention.
2. *"When a word appears in the quotation in both singular and plural form, we always prefer the (shorter) singular form". However,*
3. *"When it appears only in plural form, we can't know if it is reasonable to use it in singular form and hence we keep it in plural form".*

- First of all, let us see what prefixes we receive when we consistently select only the singular form (we will ignore the term שר מלכות because it is too long):

1. כוכב. 2. מזל. 3. מלאך. 4. שר. 5. מלך. 6. אדון. 7. קצין. 8. שוטר. 9. מושל. 10. משרת. 11. עליון. 12. עיר. 13. קדיש. 14. אלה [11]. 15. אלוהות. 16. אל.

- As it turns out, not only does this entire list fail in the permutation test in *WGP*, but the set of four prefixes, 4, 5, 12, and 14, BMS's arbitrarily chosen quartet, fails as well.
- This is why BMS had to come up with "guideline" number three. Using this guideline we receive the set:

1. כוכב. 2. מזל. 3. מלאכי. 4. שר. 5. מלכי. 6. אדון. 7. קצין. 8. שוטר. 9. מושל. 10. משרת. 11. עליון. 12. עיר. 13. קדיש. 14. אלהי. 15. אלוהות. 16. אל.

- Unfortunately this set also fails. The third guideline only helps with regard to the four prefixes, 4, 5, 12 and 14, for which BMS are trying to contrive their story. By replacing the singular form מלך with the plural מלכי and (according to

[11]) by replacing אלה with אלהי, we finally receive the desired set – i.e., the one that succeeds.

- However, the arbitrary choice of prefixes 4, 5, 12, and 14, out of a total of 15 prefixes, is made away from the eyes of the reader and no explanation is given for it.
- Guideline number three is ridiculous enough in its own right. For example, there is absolutely no doubt that מלך ("king of") is a perfectly acceptable prefix. Since according to Nachmanides שרים = מלכים, it follows that שר = מלך. Indeed the word מלך used as a prefix appears in Scripture – מלך פרס ("king of Persia"). In fact it appears in the very same chapter of Daniel from which Nachmanides took the expression מלכי פרס ("kings of Persia"). Therefore when BMS say that, "*we can't know*" whether the singular form מלך is acceptable as a prefix they are speaking nonsense.

5. The best quartet:

In this section we will describe the measurements mentioned in Chapter V, which were conceived and executed by Yosef Beremez.

1. How the "Best Quartet" is determined for any text T using **P_1**:

a. We will accept BMS's list of prefixes k_i, where $1 \leq i \leq 136$.

b. Every prefix k_i defines a set K_i of 68 word pairs (X, prefix X), where X is one of the names from the "Table of Nations" (see Appendix to Part I).

c. When we calculate the values of *c(w,w')* for the pairs in set K_i, we receive a set $K_i(T)$, in which there are $n(K_i,T)$ values of *c(w,w')*, where $n(K_i,T)$ is a number between 0 and 68.

d. We then calculate the value of P_1 for the set $K_i(T)$.

e. We rank the sets $K_i(T)$ in ascending order of P_1 values, so that $K_1(T)$ has the lowest value of P_1 (the set of word pairs for which $K_1(T)$ is its set of *c(w,w')* values, will be denoted K_1).

f. The "Best Quartet", Q(T), is defined as the union of the four sets: $K_1(T)$-$K_4(T)$.

g. To measure the significance of the statistic $P_1(T)$ for the quartet Q(T), we perform the randomization PRWL on the set of pairs Q, which is the union of the four sets K_1-K_4.

- Using this procedure on the book of Genesis we receive a P_1 value for the quartet Q(G) of: $\mathbf{P_1(G)=1.65X10^{-9}}$.
- Using the randomization PRWL with 10,000,000,000 permutations we receive a significance of: $\mathbf{r_1(G)=4.0X10^{-10}}$.
- Regarding the text *WGP* the value of P_1 for the quartet Q(WGP) was: $\mathbf{P_1(WGP)=4.23X10^{-5}}$.
- Using the randomization PRWL with 100,000,000 permutations we receive a significance of: $\mathbf{r_1(WGP)=6.16X10^{-6}}$.

2. Evaluating (through simulation) the probability of receiving a significance of **$r_1(T)$**:

Because of the enormous difference between the values of **$r_1(G)$** and **$r_1(WGP)$** we are compelled to evaluate the probability of receiving such a low value of **$r_1(T)$**. There is no easy way to make such a calculation. One cannot, for example, simply determine how many quartets can be drawn from a list of 136 items, because these quartets are not independent. Therefore we used a simulation as follows:

a) In place of the set $K_i(T)$ in step c. above, we will build a set $K'_i(T)$ in which there are $n(K_i,T)$ random numbers chosen by lottery from the segment (0,1] (this is the range of values of *c(w,w')*), produced by a random function in the computer.
b) We then carry out steps d-f as above, using $K'_i(T)$ in place of $K_i(T)$.
c) In this way we receive the set Q'(T), for which we calculate the value $P'_1(T)$. (Step g. above, which was needed to guard against the effects of any possible dependence between the values of *c(w,w')*, is unnecessary when random numbers are involved.)
d) We repeat the series of lotteries and calculations in steps a)-c) numerous times, and examine the ranking of $\mathbf{r_1(T)}$ among the set of values of $P'_1(T)$.

3. Results and Conclusions:

- We performed 1,000,000 lotteries for Genesis and calculated 1,000,000 values of $P'_1(G)$. The ranking of $\mathbf{r_1(G)}$ among this set of values was 420.
  Hence, the probability that $\mathbf{r_1(G)}$ is so small is **p=0.00042.**
- We also performed 1,000,000 lotteries for *WGP*, and calculated 1,000,000 values of $P'_1(WGP)$. The ranking of $\mathbf{r_1(WGP)}$ among this set of values was 353,949.
  Hence, the probability that $\mathbf{r_1(WGP)}$ is so small is **p=0.354.**

Acknowledgements:

I would like to thank Yoav Rosenberg and Yaakov Rosenberg for preparing the programs used in making these measurements. I'm grateful to Yosef Beremez for contributing to both parts of this article with his creative ideas and valuable work. I wish to express my thanks to Dr. S. Srebrenik for helpful discussions and suggestions.

Bibliography and Notes:

1. D. Bar-Natan, B. McKay & S. Sternberg, *On the Witztum-Rips-Rosenberg Sample* of *Nations*, preprint, March 11, 1998, the Introduction to their article.
2. Section 3.4.
3. See Table 4 in the Appendix to their article.
4. Section 7.1-2.
5. M. Bar-Hillel, D. Bar-Natan, B. McKay; *One Can Skip in* War and Peace *Too*, Galileo, No. 25 (1997), pp. 52-57.
6. See Document 4 at http://www.torahcode.co.il/english/pdf_files/docum4e.pdf.
7. M. Bar-Hillel, D. Bar-Natan, B. McKay; *The Torah Codes: Puzzle and Solution*, Chance, vol. 11, No. 2 (1998), pp. 13-19.
8. See their comment about Table 4, and at the end of Chapter 4.
9. Copies of the programs were deposited with editorial board of *Statistical Science* at the time, and they also received a copy.
10. In section 3.2.
11. See Table 5, number 5.