Bs"d

A Refutation Refuted
or: How the List of Famous Rabbis
Failed in War and Peace

doron Witztum

Summary:
In our paper, "Equidistant Letter Sequences (ELSs) in the Book of Genesis", published in Statistical Science (Vol. 9 No. 3, Aug. '94), we showed that expressions have been intentionally encoded in the form of ELSs in the book of Genesis. We accomplished this by measuring the tendency of the names of famous Rabbinic personalities to converge with their dates of birth and/or death (day and month). This analysis revealed such a strong tendency that the probability of it having occurred by chance is extremely small. According to one measure of probability the chances are only four in a million.

On the other hand, an article appeared on the Internet on September 20, 1997, under the title "Equidistant Letter Sequences in Tolstoy's War and Peace," in which Dr. Dror Bar-Natan and Dr. Brendan McKay claim to have found "the same phenomenon" described in our paper in the novel War and Peace.

Obviously the authors do not mean that there really are expressions encoded in the form of ELSs in War and Peace. Their real claim is that there are no such intentionally encoded expressions in Genesis.

Dr. Dror Bar-Natan and Dr. Brendan McKay are in essence claiming that the results of our research were obtained through manipulation - that is, by taking advantage of the latitude they claim exists in the rules we used, or by deviating from the rules altogether. Their claim is based on an experiment they carried out, the results of which they posted on the Internet, in which they were able to produce a false "success" in War and Peace through admitted manipulation within the rules mentioned above, or by breaking them "to the same extent" that we supposedly did.

In our response we show that:

1. The "success" in War and Peace was produced entirely by breaking the rules, so that the results are completely without significance. (They, too, admit that without rules one can produce whatever result one wishes).
2. They did not succeed in demonstrating a single instance in which we broke the rules while carrying out the original research.

Paradoxically, their failure to produce an artificial success within the framework of the established rules is strong evidence against their claim that within those rules there is still enough freedom to produce a comparable level of significance in War and Peace.

Furthermore, we shall see that the very points they raise serve to substantiate the integrity of our research, in particular by demonstrating that the lists of names and dates were compiled objectively and a priori.

General Overview:

In our paper, mentioned above, we showed that expressions have been intentionally encoded in the form of ELSs in the book of Genesis. We did this by measuring the tendency of the names of famous rabbinic personalities to converge with their dates of birth and/or death (day and month). This analysis revealed so strong a tendency that the probability of it having occurred by chance is very minute. One measure indicates a probability of only four in a million.

The list of names and dates was prepared in advance, following an objective procedure. The names and appellations of the rabbis were determined by an expert in bibliography, Professor Shelomoh Zalman Havlin, who was head of the Department of Information Studies, Bibliography and Librarianship at Bar Ilan University at the time the list was prepared.

[A detailed report on the principles and rules he used can be found appended here in Document 1 (which we will henceforth refer to as simply as "the Report"). A chronology of the steps and stages which preceded the publication of the paper in Statistical Science is presented in Document 2: "Bar Hillel and Bar Natan Ask - Witztum and Rips Respond".]

In the intervening years many more experiments have been carried out which point to the existence of intentionally encoded expressions in ELSs in the book of Genesis (a number of which will be presented in my new book, currently being prepared for publication). Among them is an important work by an American researcher whose expertise is in deciphering codes for the U.S. Defence Department. All of these researches were carried out using objectively compiled, a priori lists of expressions (lists which, incidentally, did not require the services of an expert consultant).

Yet in the article "Equidistant Letter Sequences in Tolstoy's War and Peace" the authors, Dr. Dror Bar-Natan and Dr. Brendan McKay (who will henceforth be referred to as: BNMK) make two central assertions concerning our paper:

Assertion 1: They criticize the rules established by Prof. Havlin. They write that "Havlin acknowledges making many mistakes in preparing the list and says that if he were to do it again, he would have done it differently."

They add to this the critique of Prof. Menachem Cohen of the Biblical Studies Department of Bar Ilan University, concerning the "objectivity and accuracy" of Prof. Havlin's list and concerning the Report.

We will respond to this criticism in full in Part I. There we will show that Prof. Havlin proceeded according to proper professional guidelines, in a manner which was objective and a priori. We will also show that the assertions of Prof. Cohen, as well as those of BNMK are groundless in every detail, and that on the contrary, by analyzing their assertions one can come to appreciate the extent to which Prof. Havlin proceeded without the slightest bias or impropriety.

To this section we append Document 3: Prof. Havlin's letter of response to Prof. Cohen.

Assertion 2: They claim further that we "still had some choice in applying their rigid procedures - enough choice to generate comparable significance levels in War and Peace." That is to say, BNMK claim that within the rules which were established before the compilation of the first list, there is enough latitude to

produce through manipulation an artificial "success" for another list in War and Peace.

They try to substantiate this claim on the basis of a list of names which they presented via the Internet on the Sept. 20, '97, in their article: "Equidistant Letter Sequences in Tolstoy's War and Peace."

It is clear that BNMK invested considerable effort in the preparation of a list of names which would succeed in *War and Peace* and fail in Genesis. It was an effort which involved searching for many sources, and a great amount of computer time to make the calculations. This list was prepared over the course of many months, and what they published was not the first version of the list. They try to justify their selections by a set of 24 assertions (section 2.1 in their article) and by the considerations mentioned in section 2.2 (ibid). They claim to have prepared their list:

"by purposefully constructing our own list of appellations, staying within the WRR-stated rules or breaking them by about as much as they did."

As we will show in Part II, their claim to have performed their manipulations within the established rules is entirely without foundation. There is no comparison whatsoever between what they did and what we did. Their entire list of modifications consists of nothing but flagrant and unjustified deviations from the rules mentioned above. Therefore their attempt to demonstrate the latitude which supposedly exists within those rules is a complete and total failure. On the contrary, it is precisely the complete and total failure of their efforts which demonstrates conclusively the spuriousness of their assertion that within the rules there remains: "enough choice to generate comparable significance levels in War and Peace.

From our responses to their list of assertions it emerges clearly that Prof. Havlin selected his list of names and appellations in a manner which was a priori and without bias.

We append to our response another document, Document 4, which demonstrates that our original research was performed in a manner which was a priori and unbiased. The proofs are built on the assertions of the critics themselves.

Part I

Introduction:
In this section we will deal with the criticisms leveled against the rules used by Prof. Havlin. BNMK write that in Prof. Havlin's report: "Havlin acknowledges making many mistakes in preparing the list and says that if he were to do it again, he would have done it differently." A closer look at the Report (Document 1) reveals quite a different picture.

In the Report no such expression is to be found. What we do find is that in the section where Prof. Havlin explains his reasons for not including in the second list certain appellations which appear in the Bar Ilan Responsa database, he indicates a number of appellations which were left out inadvertently, or for which he could no longer recall the reason they were omitted.

We tallied these omissions and found that in all only 10 of the omitted names should have been on the list. That is to say, only ten of the omitted names were between 5 and 8 letters long. Of the ten, three do not appear as ELSs in Genesis at all.

We decided to investigate what would have happened if the remaining seven names had been included in the original list. Here are the results (recall that in the original experiment the statistics P1 and P2 served as the measures of probability. This is how they were presented in the "Blue Preprint" - the permutation experiment was suggested at a later date):

Originally, the best result was:

$$P2 = 0.00000000201$$

If we add to the list the seven names which were omitted, we receive:

$$P'2 = 0.000000000101$$

In other words, the results improve by a factor of 20!

This should make it perfectly clear that Prof. Havlin did not omit these names in order to improve the result. Nevertheless, BNMK may have intended that it would be more proper to evaluate the statistical significance (using the permutation test) for Prof. Havlin's emended list. In response to this challenge we performed the permutation test with the addition of the seven names. In an experiment in which we ran 100,000,000 permutations, and P4 came in eighteenth place, that means that the probability is less than 1/5,500,000!

Now let us consider the criticisms of Prof. Menachem Cohen. He raises two issues:

1.   That "the principles according to which Prof. Havlin chose the names and appellations are completely arbitrary, and for every paragraph (in his report) one could have proposed alternative principles, which would not have been inferior, and in some instances would have been superior to the principles presented."

2.   That "the selection process itself is not consistent, even in light of those very principles, and is filled with contradictions."

Response to Criticism 1:

A.   We recommend that the reader examine Prof. Havlin's report (Document 1). By examining the Report the reader will come to appreciate the challenge which faces a bibliographer attempting to compile such a list. He will discover what the difficulties are, what factors he needs to take into account, and what the reasonable solutions are. We also recommend reading Prof. Havlin's letter responding to Prof. Cohen's criticisms (Document 3). The following discussion is based also on this letter.
Prof. Cohen's criticism that the principles laid down by Prof. Havlin are arbitrary is not true, and we will see as we go on that Cohen's words are themselves "filled with contradictions" (to use his own language).

B.   Nevertheless, it should be made clear at the outset that the supposed "arbitrariness" of the principles is, in fact, totally irrelevant:

B1.   Let us suppose, for the sake of argument, that the person compiling the list

chose his rules on a totally arbitrary basis. Or let us even suppose that he operated without any guidelines at all, writing the names and appellations according to his personal store of knowledge, without any rules whatsoever. From the standpoint of the experiment performed, this would in no way impugn the results, as long as the compiler had no prior knowledge of which names or appellations would be "successful."

Therefore, Prof. Cohen's real criticism is not against this or that arbitrary choice. Rather he is accusing Prof. Havlin of "cooking up" the rules to ensure that the names and appellations chosen by means of them would lead to the success of the experiment, based on the prior knowledge which had been made available to him.

In fact, Prof. Havlin operated without any knowledge of which names or appellations would "succeed" in the experiment and which would "fail." He was even lacking basic information concerning the exact nature of the phenomenon under investigation, and certainly about the procedure for measuring the results. (In fact, to this day Prof. does not know which names "succeeded" and which "failed," with the exception of those he learned about from this document).

Besides the absurdity of suspecting Prof. Havlin of having conspired with the authors of the paper, it will also become clear that the examples upon which Prof. Cohen bases his accusation lead to a conclusion directly opposite to the one he is trying to draw. Not only do they not call into question Prof. Havlin's integrity, they actually attest to it!

B2.   Prof. Cohen's criticism is particularly irrelevant regarding the second list, because the principles were established and employed in the compilation of the first list. Even if one were to suppose that they were chosen arbitrarily, nevertheless, they were established long before the reviewers requested of the researchers that they perform a second experiment, for the sake of which the second list was compiled. (See Document 2, Response 1 concerning the chronology of our original investigation).

C.   From all of the above it should be clear that Prof. Cohen's criticism is irrelevant. We could stop right here. But the truth is that besides being irrelevant, it is also false.

C1.   Prof. Cohen claims that "Havlin's decision to choose only those appellations which are pronounced has no scientific validity, beyond the chooser's own whim."

This assertion is very strange. The very essence of a name is that it is pronounced, as the verse in Genesis (2:20) states: "And the man **called** names, etc." A name is something that one "calls." Among the acronyms and abbreviations (which are used a great deal in Torah literature) there are those which are pronounced, and those which are simply a shorthand form of writing. (See also the Report, "Professional Judgment" sec. (a)). Therefore there was really no choice to be made here at all. After all, our subject was appellations of Torah scholars. Abbreviations which are not pronounced are not appellations.

The truth is that this complaint of Prof. Cohen contradicts his own words in a previous paragraph of his letter, in which he states that "these appellations evolved for the most part through the framework of the written literature; most of them are complete acronyms (like "Rambam"), or partial acronyms (R"Y Caro), or the names of written works (like "Beit Yosef"). In the course of time some of them

departed from the literary framework and became <u>pronounced</u> names, which at times even took the place of the original name."

In other words, Cohen also admits that there is a category of names which are pronounced, which have special status, to the point that they are capable of displacing the original name. For example, Rambam is such an appellation.

Clearly, then, this principle is not a convenient invention of Prof. Havlin.

C2. Prof. Cohen cites as an example of one of Prof. Havlin's "arbitrary principles" the fact that when an appellation is associated with more than one personality, Prof. Havlin identifies it only with the personality who was most famous, or who belonged to the period of the Rishonim (the early sages). Cohen asks:

"Why is it not possible for two people to merit being called by the same appellation if it is the one which was conferred upon them by the people of their generation and/or by succeeding generations? Furthermore, Havlin does not apply the same principle with regard to the original names of those sages. It is possible to perform a matching of names with birthdates even if four or five scholars were called by the same name! Is there any fundamental difference between the two categories?"

The answer is to be found in the words of Prof. Cohen himself. In the first paragraph of his criticism he distinguishes between a person's given name and his appellations: "The soundest principle is, in the natural course of the matter, to use the name which was given to the man at birth. It is difficult to argue with its validity." On the other hand, regarding the various appellations, "which sometimes came about and developed over the course of hundreds of years and in different locations," their status is less firm, and even variable.

Therefore, a man's given name clearly belongs to him even if someone else bore the same name. We should not deprive him of it.

On the other hand, an appellation like "Maharsha" is another matter. When can an appellation such as this be said to belong to such-and-such a personality? Is it sufficient that one of his students referred to him in this way? Or should it only be considered one of his appellations when his entire congregation called him "Maharsha"? Or perhaps only if everyone referred to him by this appellation? Should we treat this as his appellation even if it was only used during his lifetime? Or perhaps only if it was still in use three generations later?

From these questions it should be clear that it is impossible to assign the appellation "Maharsha" to anyone who was ever referred to in this way. That would border on the absurd. Prof. Havlin adopted the simplest solution: He reserved the appellation for the most important and well known sage with whom it is associated. The concept of "most important" includes the accepted distinction between the Rishonim (the "early" sages) and the Acharonim (the "later" sages). (See his Report, "Professional Judgment" sec. (b)).

Any other solution (for example, criteria of dissemination and duration) would have necessarily been <u>more</u> complicated and artificial.

Bear in mind that Prof. Havlin established this rule before the preparation of the first list. It turns out that the only instance in which this principle had to be applied in practice, in the first list, was with regard to the acronym "Rivash" for R. Yisrael Baal Shem Tov. Havlin did not include this appellation because it is too intimately associated with one of the Rishonim, R. Yitzchak bar Sheshet.

It so happens that if Prof. Havlin had acted according to Prof. Cohen's

recommendation and used the appellation "Rivash" for R. Yisrael Baal Shem Tov, the results would have **improved**:

> The results for the first list were:
> P1 = 0.000000001334   and   P2 = 0.00000000145.
>
> With the addition of "Rivash" the results were:
> P'1 = 0.000000000412   and   P'2 = 0.00000000117.

That is, the best result improved by a factor of 3.24!

C3.   Prof. Cohen cites another example of what he considers an arbitrary rule, relating to what he calls the "rules of Hebrew grammar" principle (see the Report, "Professional Judgment" sec. (e)). In keeping with this principle, Prof. Havlin rejected using the definite article before combinations of words such as: הבית יוסף, in which the article ה specifies the expression בית יוסף. Prof. Cohen claims that "one cannot ignore with a wave of the hand" a usage which exists, just because it deviates from proper grammar. He insists that the matter requires an "authoritative clarification", and that "expert linguists would call into question Havlin's decision in this matter."

In fact, Prof. Havlin did not ignore "with a wave of the hand" any existing usage. In fact, he sought the expert opinion of the linguist Yaakov Auerbach, of blessed memory. He came down firmly on the side of rejecting ungrammatical expressions such as הבית יוסף, particularly in light of our intention to look for their appearance as ELSs in the Torah. If ELSs of expressions do appear by design in Genesis, they will certainly not be written with grammatical errors! Therefore it is preposterous to suggest that we look for mistaken usages.

Even Prof. Cohen acknowledges that there is room for doubt whether an expression like הבית יוסף should be considered a mistake, as the rules of proper grammar indicate, or whether one should take common usage into consideration and stretch the rules of grammar accordingly. If the list Prof. Havlin presented to the researchers had included doubtful expressions, which might simply be mistakes (which according to the research hypothesis there would be no point in looking for in the Torah), could he have been said to have fulfilled his assignment?

As a result of this criticism of Prof. Cohen we can better appreciate not only of Prof. Havlin's professional judgment, but also his reliability - because out of all the expressions of the form הבית יוסף pertaining to the first list, not one appears as ELSs in Genesis. Therefore this   principle had no effect whatsoever on the results!

If Prof. Havlin had had any advanced information about the appearance of these names in equidistant letters he surely would have avoided establishing such a principle. After all, why should he open himself up to a situation where "expert linguists would call into question" his decisions?

Response to Criticism 2:

In this criticism Prof. Cohen tries to demonstrate a "lack of consistency" in the application of Prof. Havlin's rules. Unfortunately, as will become clear immediately, Cohen's criticisms emanate from his lack of familiarity with the subject at hand. He is an expert on the Biblical text, not an expert in bibliography.

A.   Prof. Cohen claims that the principle requiring pronounced expressions is not consistently applied and that not all unpronounced expressions were rejected. As examples he cites the appellations מהרימ״ט, מהרמ״ז, א״ח הע״יר.

Despite his assertion to the contrary, all of these expressions **are ronounced**! The proof he cites to support his contention is particularly absurd. He says: "Undoubtedly if one were to ask the average yeshiva student to explain these acronyms he would not know what you were talking about."
In fact, the "average yeshiva student" has never even heard of the personalities Prof. Cohen mentioned - not R Yosef of Trani, not R. Moshe Zacut, and not R. Immanuel Hai Ricchi - therefore it is indeed highly likely that he has never encountered their acronyms. Does that prove that they are not "pronounced" by those who are familiar with them?

B.   Prof. Cohen also claims that the principle of "rules of Hebrew grammar" is not applied consistently, and that "the material is filled with contradictions":

B1.   He asks: "What is the difference between הר״יי עמדין or הר״יי טראני (in which the article specifies a combination of words), which were included in the list, versus הבית יוסף which was not included?"

The answer is simple (and it is astounding that Prof. Cohen could make such an obvious error): הר״יי עמדין is not ר״יי עמדין with the definite article before it, it is an acronym which stands for הַרב רַבי יַעקב עמדין, just as הרמב״ם is not רמב״ם with the definite article, but an acronym which stands for הַרב רַבי מַשה בן מַימון.
See for example the Even Shushan Dictionary, in the section on acronyms, where he notes that רי״ף stands for רַבי יַצחק אלפַסי , whereas הרי״ף is not רי״ף with the definite article, but an acronym for הַרב רַבי יַצחק אלפַסי . (By the way, the abbreviation הר״יי טראני mentioned by Cohen does not appear on the list at all).

B2.   Prof. Cohen found another "contradiction" in the application of the " rules of Hebrew grammar":

"Why does Havlin not reject the many appellations of scholars in which a ה has been added to the beginning (like: המהר״ם, הרמב״ם, etc.)?" After all, he protests, the definite article is superfluous before a proper noun!

As we have already explained at length, הרמב״ם is not רמב״ם with the definite article before it.
Concerning המהר״ם, which is a generic acronym for scholars whose first names begin with the letter מ - for instance מַאיר: מַורנו הַרב רַבי מַאיר- it is sometimes necessary to add the definite article when referring to a specific מהר״ם. Similarly, מהר״ל is a generic acronym for rabbis with the name of ליב or ליוא: רַבי מַורנו לַיב (לַיוא). Among the numerous "Maharal's" was the famous Maharal of Prague, who is referred to with the definite article - המהר״ל - to distinguish him from the others.

C.   Prof. Cohen attacks what he calls "the principle of using only one appellation" (see the Report, "Professional Judgment" sec. (b)). He attacked this principle already in his first criticism (C2, see our **response** there).

C1. Here he adds that this principle can lead to "internal contradictions in the process of decision: How are we to act when the Acharon (the later sage) is more famous than the Rishon (the early sage)?" For example, if we were to ask the average yeshiva student who the "Maharam" was, he would undoubtedly reply "the Maharam of Lublin" (an Acharon), and not the Maharam of Rottenberg (a Rishon).

Here again Prof. Cohen is mistaken. There is no doubt whatsoever that the Maharam of Rottenberg is more important and better known than the Maharam of Lublin. Once again his reference to "the average yeshiva student" is baseless. Prof. Cohen did not even bother to ask the opinion of "the average yeshiva student" before speaking in his name. (Besides which, as we have already mentioned the "average yeshiva student" certainly does not have the expertise to make judgments such as these).

C2. Prof. Cohen expresses his astonishment at the absence of Maharam Lublin from both lists. From here he concludes: "I did not check the length of his columns in the Encyclopedia of Great Men in Israel, but the fact that he was omitted testifies to the [un]suitability of this tool for establishing the list."

Prof. Cohen is in error here on two accounts:
He did not bother to check, and he did not try to understand. If he had he would have discovered immediately that Moharam Lublin was omitted for the very simple reason that his entry does not list a date of birth or a date of death, whereas our lists consisted only of those personalities for whom a date of birth and/or death were supplied.
But Cohen makes a more fundamental error. The encyclopedia was meant to serve as an objective source from which to draw our list, nothing more than that. This does not imply that the list compiled on the basis of it includes every great scholar, nor was this the purpose in using it. All that was needed for the sake of the experiment was a list of famous scholars, which was selected on the basis of objective criterion (for example: "all scholars whose entry is longer than three columns, and for whom a date of birth or death (including day and month) is indicated").

D. Next Prof. Cohen attacks the "principle of selecting among similar variants of a scholar's appellations" (the Report, "Professional Judgment" sec. (f)):

D1. He asks: "Why should we reject one appellation in favor of another, if they were both in fact used by scholars?"

The answer to this question is related to the answer to Criticism 1, sec. C2. There we showed that even Prof. Cohen agrees that in comparison with given names, the status of appellations is less firm and even variable. Therefore, if there exist several similar variants of an appellation, it is sounder to use the most common and accepted variant. Here also it must be emphasized that this principle was established before the first list was prepared.
When we read Prof. Cohen's criticism, it occurred to us to investigate what would have happened to the results of the first list if Prof. Havlin had not

established this rule: That is to say, if he had included all related variants of the appellations, as well. As it turns out, the results would have improved!

The results for the first list were:
$P1 = 0.000000001334$ and $P2 = 0.00000000145$

With the addition of related variants the results are:
$P'1 = 0.000000000262$ and $P'2 = 0.00000000109$

In other words, the best result would have improved by a factor of more than 5!