

BS"D, Cheshvan 5762 (Nov. 01)

Smoke Screen

Concerning McKay's Response to our Article

"Concerning the Choices of Dates for WRR's Rabbis Samples.
Part A: Direct Optimization."

By Doron Witztum

Introduction:

Our article [1], "Concerning the Choices of Dates for WRR's Rabbis Samples" (Part A), scrutinizes McKay *et al*'s claim that WRR directly optimized the results, by exploiting "beneficial" choices pertaining to the dates. Careful examination of all the choices indicates the opposite: It shows WRR's perfect integrity. Alternative choices, based on McKay *et al*'s suggestions, would have yielded **better results** — sometimes by a factor of 2 or 3, sometimes by a factor of 10 or 100, and sometimes by a factor of **tens of thousands**. All this starkly contradicts both McKay *et al*'s report and the impression created by their article. These results, which constitute *direct* evidence for WRR's integrity, are far more impressive and significant than the *indirect* evidence sought for by McKay *et al*, in order to prove that WRR "cooked" their results.

How did McKay's reply treat this solid evidence and its conclusions?

The answer would surprise only people who are still unfamiliar with McKay's previous replies. It transpires that his reply [2], "The Choice of Date Forms – a reply to Doron Witztum," **is actually no reply at all**. Instead, it's a failed attempt to lay a smokescreen, specifically by constructing new "straw men".

How does McKay's reply tackle the direct evidence which indicates that WRR acted honestly?

In the introduction to our article [1] we wrote:

In this section we will scrutinize MBBK's claim that WRR directly optimized the results by exploiting "beneficial" choices pertaining to the dates.

Concerning direct optimization, remember that originally P1 and P2 were the sole statistics used to measure the success of L1 and L2. Therefore, any optimization of dates must have been in relation to P1 or P2, or, more probably, in relation to $\text{Min}(P1, P2)$. Therefore, it is most sensible to examine the situation with these statistics.

Accordingly, we brought the results of many experiments which show that WRR acted honestly in the choices indicated by McKay *et al*.

McKay's reply tries to evade this evidence in two ways:

1. It tries to cancel the significance of our results with the following "maneuver": Claiming false assertions, it attempts to negate the legitimacy of using measures P1 and $\text{min}(P1, P2)$. (This is done in the section named "The choice of success measure").
2. It draws one's attention from the meaning of our results – that WRR acted honestly – to a totally different subject (in the section titled "Variations that add data").

Let us deal briefly with these two points.

1. McKay claims that the use of $P1$ and $\min(P1, P2)$ was invented only **after** the publicizing of our original experiments on the rabbis lists; thus he wishes to destroy the legitimacy of our using these criteria in [1].

Our reply:

a. In our first preprint [3] we defined the following measure, which we will denote here as $P'1$. Out of the n results (for "name–date" pairs) of the whole sample, we counted the number of results whose values were smaller or equal to $p=0.2$. We denoted this number by S . The probability of getting such a big S can be calculated directly (under certain assumptions) using the binomial distribution. Because this calculation is quite complicated, we did it through the normal approximation to the binomial distribution: We counted how many standard deviations there were between S and the expected value. This number is $P'1$. In a table, we denoted the value of $P'1$ together with the values of n , p and S , so that the reader could calculate the odds of S being so great: Either by translating the value of $P'1$ into probabilities, or by a direct calculation using the binomial distribution.

For example: For the first sample we gave the following data:

Total number of pairs: $n=152$.

Number of results whose value was less than or equal to $p=0.2$: $S=63$.

Number of results whose value was less than or equal to $p=0.2$, which would be randomly expected: $E=np=30.4$.

The standard deviation: $\sigma = \sqrt{np(1-p)}=4.93$.

Number of standard deviations: $P'1=(S-E)/\sigma=6.61$.

Translating the value of $P'1$ into probability gives: 1.92×10^{-11} .

On the other hand, a direct calculation of the probability, using the binomial distribution gives the value of $P1$. In our example: $P1=1.33 \times 10^{-9}$. This is the accurate value.

After our two experiments were published in our second preprint [4], a software which directly computes $P1$ was prepared for us, and we no longer needed to use the approximation $P'1$.

McKay's claim is that in the computations involved in [1], I should have used $P'1$ rather than $P1$.

But it should be noticed, that when $P1$ gets smaller – so does $P'1$, and vice versa: when $P1$ get bigger – so does $P'1$. **Therefore, whenever the choices pertaining to dates lead to an improvement of $P1$ – $P'1$ will also be improved, and vice versa.**

Moreover, new computations reveal that the conclusions arrived at in [1] **remain valid** even when $P1$ and $\min(P1, P2)$ are replaced by $P'1$ and $\min(P'1, P2)$, respectively.

Thus, using this claim, McKay failed to invalidate the conclusions of [1]. On the other hand, McKay's claim destroys one of the central assertions of ...McKay *et al*, an assertion which relies entirely upon replacing $P'1$ with $P1$. Namely, McKay *et al* used $P1$ in their computations instead of $P'1$. If we correct this mistake (according to

McKay's present claim), their whole assertion will collapse! (See in the Appendix, A).

b. Concerning $\min(P1, P2)$:

McKay himself writes that we considered the two measures, $P1$ and $P2$, as probabilities. So, obviously, if we give two results, we would expect the reader to be impressed by the better one: This is the mathematical logic of presenting two results. Perhaps McKay doesn't understand this simple rationale? Don't worry – he does! Elsewhere, he even uses it himself to support certain of his own claims (see in the Appendix, B).

Conclusion: McKay's assertions fail to save him from the clear conclusions of [1].

2. In the section headed "Variations that add data," McKay's reply tries to distract attention from the true meaning of our results in [1]: That WRR acted honestly. McKay writes:

“Witztum presents a number of examples where adding additional data to his experiment (such as additional date forms) improves the value of $P1$ and $P2$. This can occasionally be justified in a study of choice making, though not always. Leaving that issue aside, we wish to record here why such experiments are irrelevant to the issue of whether the codes are genuine.”

McKay then proceeds to present the results of an experiment, which shows that $P1$ and $P2$ do not represent statistical significance.

Our reply:

We do not claim that $P1$ and $P2$ represent the statistical significance. $P1$ and $P2$ are the overall measures, used as a measure of success in the original experiment. To calculate the statistical significance of $P1$ and $P2$, a permutation test was devised and conducted more than two years later. All this is explained in our paper in *Statistical Science* [3].

Our article [1] also never claims that the improvements in the results of $P1$ and $P2$, or of $\min(P1, P2)$, indicate improvement of the significance of the overall result.

So what did we prove in [1]?

In [1] we examined and proved that WRR's choice of data was perfectly honest. (It's worthwhile to see the many results there that prove this). There was no optimization. This is true even where, according to McKay *et al*, this could have easily been done. We did this by examining the various changes produced in the values of $P1$ and $P2$ (which were the yardsticks of the original experiment's success) by the various alternative choices. Time after time, it was demonstrated that WRR **did nothing** to choose the more "beneficial" possibilities.

It should be noted that WRR's experiment included two components:

- a. Preparation of the data.
- b. Correct statistical measurement of the significance.

The results of the experiment depend on the correct performance of these two components.

- Concerning component *b*, it is clear that the significance must be measured with the randomization test, as publicized in our article in *Statistical Science* [3].

- Concerning component *a*, the crucial question is whether the data was prepared honestly. Indeed, the evidence brought in [1] is intended solely to demonstrate this point. It's obvious that the examination of component *a* must be conducted using the same measures of success that existed when the data was prepared. Since we proved in [1] that component *a* was done honestly, we thus proved that WRR's result is correct. This proved that the Genesis Code is genuine.

Nevertheless, McKay's reply tries to create the misleading impression that we treated P1 and P2 as statistical significance in [1], and that our conclusions there are based on this error. This is nothing but an attempted smokescreen.

McKay's "straw man" concerning the date forms

The overwhelming evidence in our article [1] that the date forms were chosen with honesty and integrity, obviously embarrassed McKay. We show there, that the natural extension of the list of date forms improves the results according to the original criteria of success. We also show that there is an improvement even if we include the pair of date forms "א' לתשרי" and "בא' לתשרי" (which McKay *et al* had suggested, even though they are not included in a natural extension). For the reader's convenience, we quote in Appendix (C) the relevant material from [1] (Sec. 6 of its Appendix).

McKay's reply ignores all this. Instead, he tries to deflect the discussion to some marginal issue. But since, even there, he has little to say, he also sets up an appropriate straw man, as we will see later.

The marginal issue discussed this time, is how to present the frequency of several date forms in Table 16 of [1].

What is Table 16?

Through table 16 (see Appendix, C) I demonstrated that the forms

- א' תשרי.
- בא' תשרי.
- א' בתשרי.
- בא' בתשרי.

are not only standard in *Encyclopedia Margalit* (EM) and *Encyclopedia Hebraica* but also standard and widespread in the Hebrew language. To this end, I conducted a survey regarding the use of the various date forms. The survey was done using a database, which, as far as **Hebrew** is concerned, is similar to these encyclopedias: The database of "modern *Halachic* authorities" (taken from the computerized Responsa database of the Bar-Ilan University). The results of this survey are listed in Table 16. The prevalence of the variant date forms was presented for the pairs of date forms I, II, and III, which were defined as follows.

- I= "א' בתשרי" + "בא' בתשרי",
- II= "א' בתשרי" + "בא' בתשרי",
- III= "א' לתשרי" + "בא' לתשרי",

The results show that the prevalence of III is far smaller than that of I or II.

The straw man's conception:

McKay's reply proceeded to set up a straw man:

1. He ascribed to me the claim that of the four standard forms *a-d*, forms *a-c* are the commonest.
2. He ascribed to me the claim that the sole criterion for using a date is its frequency of use:
 “His case is that the common forms should be used and the uncommon forms should not be”.

In particular, he ascribed to me the position, that the non-use of the forms "א' לתשרי" and "בא' לתשרי" was based purely on their frequency in the computerized Responsa database of the Bar-Ilan University.

The next stage included the invention of the following fiction: That Table 16 is the result of a deliberate deception (what otherwise?) on my part. McKay's reply utilizes the straw man to show, that the method of presenting **pairs** of dates in the table, and not single dates, combined with the fact that I only use the database of "modern *Halachic* authorities" (and not the entire Responsa database) is meant to conceal:

3. That forms *a-c* are not the most prevalent, and that I therefore acted contrary to the established "criterion" by not using form *d* in the original experiment.
4. That I acted contrary to my own criterion, when, in the process of extending the list of date forms (done in the Appendix of [1], section 6 - see here, section C of the Appendix), I added the pair of forms "א' דתשרי" and "בא' דתשרי".

The straw man's death:

McKay's straw man is a combination of a selective and truncated quotation, followed by misleading interpretation. By dint of simply examining the facts – the straw man vaporizes and vanishes.

In the Appendix (D) I prove that the two claims which McKay ascribes to me (1 and 2 above) are figments of the imagination; therefore “facts” 3 and 4 **are false**. Consequently, McKay's conclusions are baseless. Consequently, there is no basis for the conclusions towards which McKay tries to drive the reader.

Let us once more remind the reader, that the whole issue of Table 16 is of marginal importance; it is merely meant to demonstrate that the prevalence of III is tiny compared to that of I and II, and nothing more. In fact, the data presented by McKay shows **the same thing**.

Pairs of date forms:

The most surprising aspect of this protracted tale is that McKay knows perfectly well that there is no "mystery" in the fact that Table 16 was presented through **pairs** of date forms. According to **his** logic – this should be the **natural course**. It is McKay *et al* who wrote [6] that the adding of date form *d* would be the most natural:

“The most obvious variation would have been to add the form akin to ‘on 1st of May’.” (Pg. 168-169)

Why do they consider this variation the most obvious?
 Let us explain:

- According to them, from among the four forms *a-d*, only **two** basic forms exist: Form *a* ("א' תשרי") and form *c* ("א' בתשרי"). Forms *b* ("בא' תשרי") and *d* ("בא' בתשרי") result from adding the letter "ב" to forms *a* and *c* (respectively). The letter "ב" (which serves as the word *in* or *on*) serves merely as preposition to link the date to other words in the text. [For example: The preposition "on" in the phrase "on the 27th of April" is not part of the date]. Therefore, **according to McKay et al**, the forms *a-b* are essentially one date form, and the forms *c-d* are also essentially one date form. This is why they think that variation *d* should be added to variations *a-c*.
- When they suggest additional date forms, "א' לתשרי" and "א' של תשרי", they do this in **pairs**: * "א' של תשרי" and * "בא' של תשרי", * "א' לתשרי" and * "בא' לתשרי". Each pair essentially represents *one* date form.

Therefore, specifically according to the logic of McKay *et al*, if the relative prevalence of date forms in a certain pool is sought for, it should be done in pairs – as we indeed did.

So why does McKay pretend to be unable to fathom the reason we did this?

An observation

In the section titled "One final observation" McKay describes his observation: He noticed that while the success of WRR's first list hinged mainly on the date form "בא' תשרי", the success of the second list hinged mainly on the form "א' תשרי". McKay regards this as a contradiction and a disproof of the Torah Codes.

But, this is nonsense.

First, McKay forgot that WRR's work was aimed at proving *the existence* of a hidden text in Genesis. Having such a proof at hand does not necessarily enable one to decode the language of the hidden text: One must first recognize the syntax and know far more about the vocabulary of the hidden text. (The parable in the introduction to our *Statistical Science* paper [3] may help to explain this point.)

Thus, McKay's claim concerning the "contradiction" is mere speculation concerning the unknown syntax of the hidden text.

The surprising point is that McKay forgot another observation which he had made.

In the previous section, "Pairs of date forms," we mentioned that according to the logic of McKay *et al*, the two date forms,

a. "א' תשרי"

b. "בא' תשרי"

are actually one basic date form. Form *b* is created from *a* by adding the preposition "ב", which merely links the date to other words in the hidden text.

If we combine both of McKay's observations, it turns out that the *same* basic date form succeeded in both lists!

Appendix

A. It will be shown here, that one of McKay *et al*'s main assertions rests entirely upon the replacement of P'1 by P1: McKay *et al* used P1 in their computations

instead of $P'1$. If we correct this mistake (according to McKay), and replace $P1$ by $P'1$ - their whole assertion will collapse!

McKay wrote a paper together with Kalai and Bar-Hillel [7], which developed the following story:

1. They noticed that the value of the measure $P2$ for the *first* list of Rabbis ($L1$) was very close to its value for the *second* list of Rabbis ($L2$), with a ratio of 1.12. The *a posteriori* probability for such event is (according to them) approximately 0.01.
2. They assert that WRR had “naive expectations” to get “similar success” in their second experiment. Therefore WRR “cooked” the data for $L2$ to get a success most similar to that of $L1$.

(See also McKay *et al*'s *Statistical Science* paper [6], section 8).

But:

1. The probability calculated by McKay *et al* is **meaningless**, since the number of potential *a posteriori* observations is so vast: Such exceptional *a posteriori* observations, with similar *a posteriori* probabilities, can be found for **every** experiment.

2. McKay *et al*'s result is erroneous.

Their result is based on the values of $P2$. Why specifically $P2$?

They assert in their paper [7] (at the beginning of section 3):

“For our purpose, the $P2$ value remains pertinent, since any bias in the data selection occurred at the time when it served as **the principal measure of significance**”. (Emphasis mine)

But this assertion has no basis in WRR's preprints. WRR merely presented the values of $P'1$ and $P2$, and never presented $P2$ as “the principal measure of significance” (see [3,4]).

McKay *et al*'s assertion could only be based mathematically on calculating $\min(P1, P2)$. In fact, **for both lists** we get: $P2 = \min(P1, P2)$.

This is the mathematical basis for their assertion that $P2$ was the “leading” measure in both experiments.

But, according to McKay's claim [2] that we should consider $P'1$ and not $P1$, we should calculate $\min(P'1, P2)$ and not $\min(P1, P2)$.

In such a case we would get **for both lists**: $P'1 = \min(P'1, P2)$.

Therefore, in both experiments, the $P'1$ value was the “leading” one. Hence, if WRR “cooked” the data for $L2$ to get a success closest to that of $L1$ (as McKay *et al* claim), they should have taken care that the **$P'1$ values** be similar. Let's see what really happened:

The “leading” value for $L1$ was $P'1 = 1.92 \times 10^{-11}$

The “leading” value for $L2$ was $P'1 = 3.87 \times 10^{-10}$.

The ratio of these values is 20.16 – which completely opposes McKay *et al*'s naive expectations!

B. McKay well understands, that by presenting *two* results, we expect the reader to be impressed by the better one. For example, in the “study of variations” described in McKay *et al*'s paper [6], they checked the measure $\min(r1-r4)$ for the *first* list ($r1, r2, r3, r4$ are the ranks in the permutation test), in spite of McKay's present claim that WRR did not define a measure of “minimum” for this *first* list!

C. An excerpt from [1] (Appendix, sec. 6)
(The references here refer to the sources in the original article [1].)

6. The choice of date forms.

Most of the dates pertaining to L1 are given in *EM* in standard forms and not specified by “special days”. Of the 37 dates in L1, 30 are given in standard forms. *EM* used four standard forms:

- a. "א' תשרי".
- b. "בא' תשרי".
- c. "א' בתשרי".
- d. "בא' בתשרי".

The linguist Ya'akov Orbach, WRR's linguistic advisor, suggested using the three standard forms a-c. We do not know his reasons, and we specifically do not know whether he examined or considered the forms used by *EM*. (Perhaps it is just a coincidence that the date forms used by Encyclopedia Hebraica for the rabbis of L1 are precisely forms a-c.).

(A) MBBK wrote concerning this:

“To write the day and the month, WRR used three forms, approximately corresponding to the English forms “May 1st,” “1st of May” and “on May 1st”. They did not use the obvious “on 1st of May,” which is frequently used by Margalioth...” (Pg. 155)

They also wrote:

“The most obvious variation would have been to add the form akin to “on 1st of May”. It gives the score [1.2, 2.2; 0.6, 16.4].” (Pgs. 168-169)

We examined MBBK's “most obvious” choice of including the fourth form, d, as well. Let us check the following choices:

1. Forms a-c (used by WRR).
2. Forms a-d.

The results are:

Choice no.	P1	P2	Min(P1-P2)
1(WRR)	1	1	1
2	<u>0.3</u>	1.2	<u>0.3</u>

Table 14

Note that the result improves contrary to the result given by MBBK! (As we proved elsewhere [3], their method of presenting results is designed to conceal results like these).

(B) MBBK had further suggestions to widen the choice of standard forms.

We must emphasize once more that the forms a–d are the most standard and widespread in Hebrew, and are used not only by *EM*, but also by Encyclopedia Hebraica and similar works. Any other form is rare compared to these and it is extremely doubtful whether it may be regarded as a choice. In any case, if MBBK were searching for additional forms, they should have been consistent and first looked for them in *EM* which they refer to at every opportunity.

(1) Here are the possibilities of expanding the list of date forms, while adhering to *EM*. For a complete picture we will start with the choice already examined in (A):

1. Forms a-c.
2. Forms a-d.

(2) MBBK already suggested dates specified by “special days” (Sec. 4 above) mentioned in *EM*. For these dates *EM* used the possessive word “של” (“shell”) and the possessive letter “ד” (“de”) to express dates. With this usage we get the following forms.

- e. “א' של תשרי”.
- f. “בא' של תשרי”.
- g. “א' דתשרי”.
- h. “בא' דתשרי”.

(Forms e-f were suggested also by MBBK.) Adding these choices to the previous ones brings us to the next choice:

3. Forms a-h.

(3) Surprisingly, MBBK suggested two *other* forms.

- i. “א' לתשרי”.
- j. “בא' לתשרי”.

These two forms are not only absent from *EM* (and Encyclopedia Hebraica), but they are also rarely used (see Table 16).

However, to complete the picture, we will also examine the following choice.

4. To take all the forms, a-j.

The results of these choices are:

Choice no.	P1	P2	Min(P1-P2)
1(WRR)	1	1	1
2	0.3	1.2	0.3
3	0.007	0.09	0.007
4	0.4	16.5	0.4

Table 15

It turns out that even adding forms i-j yields a result **2.5 times better** than WRR’s original!

Conclusion: The results speak for themselves: Beyond any doubt, WRR acted with perfect integrity in their choice of date forms!

(C) Concerning the frequency of the forms i-j:

At the beginning of the controversy [14] we wrote, concerning their suggestion to use form i:

“This is a nonstandard form of referring to a date. For example, both Margalioth’s encyclopedia, as well as the Encyclopedia Hebraica use the forms we used, and not this form. It is clear that the forms we used are the most widely used forms. We conducted a survey regarding the use of the

various forms, using the computerized responsa database of Bar Ilan University. Here are the results for a pool of modern Halachic authorities:

We will categorize the forms as follows:

Form I is the pair of forms: "א' תשרי" + "בא' תשרי" ("בא' תשרי" = in "א' תשרי")

Form II is the pair of forms: "א' בתשרי" + "בא' בתשרי" ("בא' בתשרי" = in "א' בתשרי")

Form III is the pair of forms: "א' לתשרי" + "בא' לתשרי" ("בא' לתשרי" = in "א' לתשרי")

The following table sums up the frequency of I, II, and III.

Month	Forms		
	I	II	III
Tishri	178	51	2
Cheshvan	364	130	1
Kislev	409	90	0
Theveth	375	108	0
Shevat	434	190	4
Adar	582	159	6
Nisan	303	126	0
Iyyar	359	82	0
Sivan	319	86	0
Tammuz	419	181	2
Av	68	263	0
Elul	286	86	0

Table 16

MBBK certainly exaggerated when they described forms III as “regular date forms”.

D. Destroying McKay's straw man concerning Table 16:

From the previous section we see, that contrary to McKay's claims:

1. I never claimed that date form *d* is less common than forms *a-c*. I just wrote as follows:

“We must emphasize once more that the forms *a-d* are the most standard and widespread in Hebrew, and are used not only by *EM*, but also by *Encyclopedia Hebraica* and similar works. Any other form is rare compared to these and it is extremely doubtful whether it may be regarded as a choice.”

Moreover, I certainly didn't explain that forms *a-c* were chosen by the linguist Yaakov Orbach o.b.m. because they were the most prevalent. Indeed, I once more explained that **I did not know** what guided him in this choice. (McKay himself quotes this statement, unaware that it contradicts the very basis of his straw man.)

Therefore, it was never necessary to prove and demonstrate in Table 16, that specifically date forms *a-c* were the most common.

2. I set up no criteria of the type that McKay tries to ascribe to me. The objective of the previous section was to find the natural extension of the list of date forms. Concerning this we wrote:

“Any other form is rare compared to these [a-d] and it is extremely doubtful whether it may be regarded as a choice. In any case, if MBBK were searching for additional forms, they should have been consistent and first looked for them in *EM* which they refer to at every opportunity.

(1) Here are the possibilities of expanding the list of date forms, while adhering to *EM*...”

And our rationale for the negation of forms "א' לתשרי" and "בא' לתשרי" was as follows:

“(3) Surprisingly, MBBK suggested two *other* forms.

- i. "א' לתשרי".
- j. "בא' לתשרי".

These two forms are not only absent from *EM* (and Encyclopedia Hebraica), but they are also rarely used (see Table 16).”

In other words, it was determined here, that all the other forms are rare compared to forms a–d, and therefore the natural extension of the list of date forms should include those forms which have the advantage of being used in the encyclopedia.

Therefore, contrary to McKay's claims:

3. We did **not** act contrary to the "criterion" when form *d* was not used in our original work.

4. I did **not** act contrary to my criterion when I indicated that the choice of the forms "א' דתשרי" and "בא' דתשרי" (the usage of the letter "ד" here is legitimate in Hebrew, and is used by the Encyclopedia Hebraica) is preferable to the date forms "א' לתשרי" and "בא' לתשרי" (the like of which are not found in the encyclopedia).

Therefore, the "explanation" McKay "cooked," concerning the method used to present date forms in Table 16 – is completely baseless.

5. McKay also asks why I used only the database of "modern *Halachic* authorities", to demonstrate the frequencies of the pairs of date forms. Here, too, the answer is simple. In order to demonstrate “that the forms a–d are the most standard and widespread in Hebrew,” and not only in the encyclopedia, I looked for a database similar to the encyclopedia in its usage of Hebrew, and found the database of "modern *Halachic* authorities" to be appropriate.

Note that even according to the database checked by McKay, the prevalence of III is much less than that of I and II, as we demonstrated in Table 16.

References

1. Witztum, D. (2001). Concerning the Choices of Dates for WRR's Rabbis Samples. Part A: Direct Optimization. Available at: http://www.torahcode.co.il/english/pdf_files/date1e.pdf.
2. McKay, B. D. (2001). “The Choice of Date Forms – a reply to Doron Witztum”. Available at <http://cs.anu.edu.au/~bdm/dilugim>.
3. Witztum, D., Rips, E. and Rosenberg, Y. (1986). Equidistant letter sequences in the Book of Genesis. Preprint.
4. Witztum, D., Rips, E. and Rosenberg, Y. (1988). Equidistant letter sequences in the Book of Genesis. Preprint.
5. Witztum, D., Rips, E. and Rosenberg, Y. (1994). Equidistant letter sequences in the Book of Genesis. Statist. Sci. 9 No. 3 429-438.

Available at: http://www.torahcode.co.il/pdf_files/pub/wrr.pdf.

6. McKay, B. D., Bar-Natan, D., Bar-Hillel, M. and Kalai, G. (1999). Solving the Bible Code puzzle. *Statist. Sci.* 14 No. 2 150-173.
7. Kalai, G., McKay, B. D. and Bar-Hillel, M. (1998). The Two Famous Rabbis Experiments: How Similar is Too Similar? A discussion paper of the Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem.