

Restoring misconceptions

The balancing effect in brain-machine interaction

There exist a number of misconceptions regarding the ideas presented in my paper “The balancing effect in brain-machine interaction” (<http://arxiv.org/abs/1602.00808>), as for instance in the role of the funnel plot of MicroPK data, or in the use of the Markovian model (that had nevertheless successfully replicated it), in the Rescaled Range Analysis of MicroPK data sequences and finally in the interpretation of the results of analysis.

Through this text certain ideas will be further clarified by directly answering some of the more relevant comments addressed. The citations here are listed in the said paper.

1. *“Actually what she says is that $p(1) = p(0) = 0.5$ and $p(1,0) = p(0,1) > 0.5$ -- I find that logically very ugly, because it is obvious that (0,1) and (1,0) can be mapped (in hard-or software) on ‘1’ and the (1,1) and (0,0) on 0 and then one has actually what Fotini keeps repeating is non existing MicroPK.”*

My response: Assuming that the comment indicates by $p(1,0)$ the probability, (or frequency), of dyad sequences of the kind: ‘10’ and likewise ‘01’, the statement $p(1,0) = p(0,1) > 0.5$ is wrong. This frequency, estimated by the Markovian model for the MicroPK database, is approximately 17%; below and not above 0.5 as this reviewer concludes. So, if “we mapped (0,1) and (1,0) on 1”, then there will be less ‘1’ than ‘0’ digits. That’s a clear imbalance of bits, not a “non-existing MicroPK” as the reviewer concludes. Yet, my paper shows quite the opposite to be the case.

The serious error seems to be the obvious confusion about what the digits ‘1’ and a ‘0’ represent in the Markovian model of MicroPK data. Each MicroPK ‘1’ involved in the funnel plot implies a ‘hit’, a successful trial. Each MicroPK ‘0’, implies a ‘failure’, a ‘miss’ and not bits generated by RNG’s. [Bösch et al 2006] have carefully converted (from ‘z-scores’, or ‘es’) all records of MicroPK tests with true RNG’s into proportion of hits, ‘pi’. The funnel plot presents the size of study, N, against this proportion of hits, ‘pi’, i.e. the proportion of 1’s.

Therefore, in the Markovian model, a dyad sequence {(0,1) or (1,0)}, i.e. ‘hit after miss’, or ‘miss after hit’, cannot be mapped by a ‘hit’, by ‘1’! Similarly, a ‘hit after hit’ and ‘miss after miss’, i.e. (1,1), (0,0), cannot be mapped by a ‘miss’, by ‘0’. As these digits are not RNG generated bits they cannot be manipulated by software or hardware. If such replacement were introduced in real MicroPK data sequences, as suggested, this would clearly be condemned as data manipulation! Neither can such mapping be implemented in the Markovian model, as equally forbidden.

2. *“In any case, I have discovered that the substance of the paper is already published ... and so fails to meet the original contribution criterion”.*

My response: This paper was written specifically in order to show that there is no discrepancy between my early test results that indicated the “balancing effect” and my recent results that indicate there is no evidence for the MicroPK hypothesis. Naturally, all related previously published results had to be invoked in this paper. As such, my paper maintains its originality.

3. *“The paper seeks to explain funnel plots of effect sizes (ES) versus study size, N, in the BSB database. The plots show 1) a convergence close to the null ($ES = 0.5$) for large N; 2) a significant increase in the dispersion of ES's; 3) an asymmetry that skews to $ES > 0.5$ for experimental data and $ES < 0.5$ for control data. The author proposes that a Markovian process (MP) (her 'gluing' effect) accounts for all of these features.”*

My response: The convergence of the funnel plot of MicroPK data is not close, but equal to 0.5. Also, the funnel plot refers to one whole database; it’s not a funnel plot for large N, or a funnel plot for small N. Therefore, the funnel plot converges to 0.5 for the whole MicroPK database. Said differently, the most representative effect size for the MicroPK database is 0.5.

The Markovian model can successfully simulate the broadening of the funnel plot and its convergence to 0.5. It’s wrong to include the asymmetry into its successes. The asymmetry in the spread of data on the funnel plot is caused by publication bias driven by the attitude of experimenters to neglect reporting

data, or to err during data collection. Once data points on the funnel plot are generated via the Markovian model, one should randomly remove some of them from sections that align with the experimenters' attitude that introduces publication bias, to account for the asymmetry.

4. *"The funnel plot features were thoroughly examined ...and debated in detail in published responses. The paper ignores the alternative explanations presented there and one would want some discussion of why they are not viable".*

My response: The reviewer of my paper implies, as it was argued in the debate¹, that the body of MicroPK data with pure RNG's be truncated into smaller parts, which are conveniently tagged according to a property of the database (e.g. the size of study), in order to separately analyse smaller parts of the database. In that sense, a selection of data will be introduced and many interpretations will be individually offered for the MicroPK hypothesis for each separate subdivision of the main database, as if it wasn't one and only hypothesis to be tested.

As it was discussed in my paper, in studies of smaller size (often generating bits at slower rate), there is a higher risk to introduce biases during the collection of data (often to satisfy the expectations of the experimenters). Those who adhere to such fragmentation of the database instinctively understand, or experience first-hand, that a stronger effect is to be expected in small size studies of the MicroPK effect. So, they emphasize the need to treat small size studies separately. Yes, small studies tend to show higher MicroPK scores, but this is not because some direct Mind-Matter Interaction manifests better in such small studies.

My paper presents the analysis of MicroPK data with pure RNG's as a whole, as the question under the microscope is only to investigate "the MicroPK effect with true Random Number Generators". Fragmentation of the database is equivalent to data manipulation and this is my answer to why such 'alternative approaches' are not viable.

5. *"In particular, the author states that she accepts clairvoyance as a psi effect but doesn't address why, say, clairvoyance à la DAT plus publication bias shouldn't offer a compelling alternative to her rather complicated mechanism".*

My response: I have suggested that "there most likely exist real psi effects, e.g. telepathy and clairvoyance" worth of investigation, which is far from stating that I have accepted them as real psi effects.

Furthermore, the reviewer prompts to adopt the following non-scientific task: to explain away a purported effect (MicroPK) by invoking another unsubstantiated effect (clairvoyance).

Finally, the application of the R/S analysis and the use of Markovian processes are not believed to be complicated by a large number of scientists.

6. *"The reasoning (as best I can follow it) goes something like this. A rescaled range analysis (RSA) on a subset of RNG data from the PEAR consortium replication finds a Hurst scaling exponent (H) greater than 0.5, and this can be taken as evidence for PK-MP correlations in the data".*

My response: Using the label "PK-MP correlations" (MP for Markovian process) that the reviewer adopts is misleading, for a couple of reasons. First and foremost, my paper shows that there is no evidence for a MicroPK effect.

Furthermore, the Rescaled Range analysis, R/S (or RSA as the reviewer labels it) does not provide evidence of a Markovian process (the 'PK-MP', according to the reviewer's tagging), but of possible long-range correlations present in the data sequences, not caused by Mind-Matter Interaction (MicroPK).

The Markovian model was only implemented to introduce a magnifying glass into the inner machinery of the MicroPK process, at the level of single trials. The model has successfully simulated the main features of the database, i.e. the broadening of data (indicating the presence of Markovian correlations between trials) and its convergence to 50% (indicating that the MicroPK hypothesis is refuted).

The collective evidence, considering the R/S analysis too, suggests that those long-range correlations present in the data sequences are introduced by “data handling”, i.e. conscious or unconscious errors during the collection and reporting of data.

7. *“Modelling the PK-MP shows that the effect can produce funnel plots with a large dispersion”.*

My response: Regarding the PK part of the PK-MP label: My analysis does not suggest that there is a PK effect. Quite to the contrary; it shows that there is no MicroPK effect.

Regarding the MP part of the same label, the Markovian model can simulate the funnel plot of a database comprised by the proportion of hits generated by a binary process. Such funnel plots can, therefore, exhibit either larger or narrower dispersion than is normally expected, or even no deviation from the normal at all (see fig. 5).

8. *“The RSA finds a highly significant H for experimental and control data. This is used to argue for PK-MP correlations between trials”.*

My response: There are two R/S analyses discussed in my paper, not only one. They were performed on two separate occasions and on different sets of data, yielding different results.

1. The first analysis was applied on the FAMMI MicroPK, control and calibration data [Pallikari, 1998; Pallikari 2001]. It indicated that weak persistent long-range correlations were identified in sequences of MicroPK data; even weaker correlations were present in control data and none at all in calibration data generated by RNG's that were previously tested for proper performance.

2. The second was applied [Pallikari, 2015] on the time-series of the MicroPK data taken from the funnel plot, i.e. on the sequences of MicroPK test records that were specifically arranged per date of publication as accurately as possible.

The second analysis indicated that persistent long-range correlations were present in the sequence of MicroPK data arranged per date of publication. In other words the effect sizes reported in MicroPK tests were not independent from one another, indicating a characteristic of mimicking in these reports.

9. *“Markovian transition probabilities of $p_{00} = p_{11} = 0.83$ are needed to reproduce the funnel plot dispersion. This is a fantastically large PK effect that would be evident in the data with simpler analyses than the RSA”.*

My response:

1. The Rescaled Range Analysis, R/S, did not produce these Markovian transition probabilities. The fitting of the Markovian model on the funnel plot of MicroPK data has produced them.

2. The Markovian transition probabilities, p_{00} and p_{11} , represent the average frequency of runs of size 2 of same bits. They are the probabilities that a ‘hit’ follows a ‘hit’ and a ‘miss’ follows a ‘miss’, in a sequence of all MicroPK records present in the meta-analysis. True, such information cannot be available for practical reasons. Yet, such high frequency is to be expected due to the longer runs of MicroPK ‘successes’ or ‘failures’ being generated.

3. Which exactly are these "simpler analyses" the reviewer refers to? And can these analyses really estimate the frequency of runs of same score of size = 2, (hit-hit & failure-failure), as in #2 above, across all MicroPK test results?

4. These transition probabilities do not indicate PsychoKinesis, PK, unlike to what the reviewer maintains.

5. An average frequency of two ‘hit’, or ‘miss’, outcomes in a row across all MicroPK experiments as high as 83% corresponds to a correlation coefficient between the adjacent MicroPK records of 66% [see table 2 in: Fotini Pallikari, *Investigating the Nature of Intangible Brain-Machine Interaction*, Journal of Social Sciences and Humanities, 1(5), 499-508, (2015)]. This is rather a moderate and not a fantastically large degree of correlation, unlike what this reviewer believes, yet surely not a ‘PK effect’.

If all experimenters carried out their MicroPK tests religiously in the same (unnatural) manner, so that the correlation coefficient between adjacent trial scores was 66% (instead of 50%) due to a fixed 83% persistence to yield a ‘hit’ after a ‘hit’ and a ‘miss’ after a ‘miss’ (instead of 50%) and their average scores were presented on a funnel plot, then this plot would share the exact same main characteristics

as the funnel plot of MicroPK data (lacking the publication bias). To account for the publication bias, some experimenters, especially of small size studies where the proportion of hits may have, by natural causes, fallen below 50%, should simply remove their data from the funnel plot. Then the observed asymmetry will be reproduced, too.

Yet, a universal mechanism of direct Mind Matter Interaction, where the MicroPK test participants affect the random process through direct mental interference, does not exist (as the database's funnel plot confirms-convergence to 50%).

In conclusion, these proclaimed by the reviewer as "high" transition probabilities stand only as an equivalent average of a model mechanism, operating deep at the level of trials, being introduced by the mishandling of MicroPK data.

10. *“However, the trial variance in all the Consortium data, including FAMMI, is at the null expectation. This entirely refutes PK-MP hypothesis, at least on a scale that would reproduce sufficient funnel plot dispersion”.*

My response: The broadening of variance that my paper treats mathematically refers to all the MicroPK test scores with pure RNG's. That is, to the unnaturally broadened dispersion of MicroPK scores from one experiment to the other in the large MicroPK database. It is the dispersion of data points in the funnel plot and not just the dispersion of trials in one experiment; i.e. dispersion of trials inside one data point on the funnel plot, such as the one referring to the consortium data.

It is well known that the consortium data have all yielded zero meanshift (refuting the MicroPK hypothesis). They may probably also have produced variance at the null expectation with respect to random data, as the reviewer asserts. Such null deviation from chance in variance, however, implies that the data are randomly distributed (see fig. 5), that there is absence of correlations in the data sequences. But, this is doubtful as my analysis [Pallikari, 1998; Pallikari 2001] has identified persistent correlations in these data sequences which should affect their variance. This reviewer seems to be well informed about the details of the consortium experimental results, but I have also got direct knowledge of them as far as the analysis of FAMMI data (of the Freiburg, IGPP branch).

The broadening of funnel plot, though, is not brought about by data in only one experiment, i.e. not due to one "star-system", but to the whole galaxy of MicroPK data. And all these data collectively exhibit broadening of variance; i.e. correlations binding data points on the funnel plot, whereas these should normally be independent from one another.

Therefore, even if there was no variance deviation from expectation in one MicroPK experiment only (which my previous analyses challenge), this does not refute the successful application of the Markovian model (which indicates increase of variance across all experiments). Moreover, the Markovian behaviour of MicroPK data is not itself PK, according to the reviewer's comment.

11. *“An explanation of the RSA consistent with the variance is that short periods of psi-hitting and psi-missing among trials causes some internal correlations that are detected by the RSA”.*

My response: The R/S analysis identifies long-range correlations, not short-range ones. These long-range correlations identify the presence of an overall trend over all MicroPK test results, one that binds the data regardless of their separation.

12. *“BSB discussed that the effect sizes decrease (more or less linearly) with publication date and there is a simple explanation for it. Most of the trend comes from early, significant studies from Schmidt's laboratory (authors Schmidt and Kelly in the database). These account for about 10% of the BSB studies, but most of the trend. With the studies removed, the RSA exponent loses most of its significance”.*

My response: The analysis of MicroPK data presented in my paper considers the database generated in the associated meta-analysis [Bösch et al, 2006] (tagged as BSB by the reviewer), as the product of very careful and honest data selection², one that provides valuable information about the MicroPK hypothesis with true RNG's.

If data are removed from this database for any reason, (e.g. accusing experimenters that they have published the wrong data, or that some studies have the unreliable effect sizes, etc.), then this is destroying/manipulating the database formed under a precise question; to investigate the hypothesis on

direct mind-matter interaction with true RNG's. One cannot apply their unjustified beliefs to dismantle a database in order to explain away their expectations. Such comment demonstrates exactly the point put forward in my paper, i.e. how easily can data manipulation be introduced in a database.

13. *"The author uses PK-MP to provide the distribution. She then claims that the positive/negative asymmetries in the experimental/control funnel plots can be explained from publication bias. But without a viable PK-MP effect, the experimental database asymmetry cannot be reproduced"*.

My response: I am not using PK-MP because the evidence provided indicates that there is no Micro-PK effect. As about the MP tag, the Markovian model successfully simulates the main characteristics of the MicroPK database: its broadened scatter and its convergence to 50%.

The database asymmetry, on the other hand, could be introduced a posteriori, if some data were removed from the already simulated database (corresponding to data that some experimenters decided not to report, thus introducing publication bias) and not because these data could already exist in the database simulated by the Markovian model. If some data points are removed from the simulated funnel plot of fig. 5c, for instance, then the asymmetry will be reproduced.

14. *"Unfortunately, the control funnel plot asymmetry is apparently an artefact of recording errors in the BSB database. In figure 4, the asymmetry is evident as a group of studies all at the same N that stretch out to the left of the plot. The 20 "control" studies all derive from a single 1979 paper (Kugel; ref ID 806 in the database). There is just one experimental study from the paper and it has a slightly positive effect size. It would be surprising if Kugel reported 20 controls for one experimental study. Typically there are fewer control studies reported in papers which are why the control database N is only a quarter of the experimental one. I strongly suspect that BSB confused control and experimental labels when creating their database. There are other instances of mislabelling in the control database: in 12 cases the observed and theoretical hit rates are inverted. If the "control" studies from this one paper are removed, there is no significant asymmetry remaining in the funnel plot."*

My response: The reviewer accuses the authors of the already published MicroPK meta-analysis [Bösch et al 2006] without providing precise and concrete evidence to support these accusations, other than to 'strongly suspect' the hypothetical misconduct.

Still, this asymmetry in the funnel plot of control data is not limited to regions of data where the reviewer focuses having size just above $N=100$, but it is also confirmed by the spread of data in other areas of the funnel plot, e.g. above $N=1000$.

Nevertheless, there is a more serious problem here. Just because the reviewer suspects erroneous data without evidence, the suggestion is put forward to remove the 'guilty' data from the database, so that existing suspicions and personal hypotheses are fulfilled. But, no data can be selected out of a carefully generated database, same as no data can be added to it.

15. *"The author claims that a statistical "balancing effect" is evident when combining the unweighted averages of control and experimental effect sizes, since these average to the null. This comparison no longer holds if the Kugel studies are mislabelled. The Kugel paper is in German and not easily accessible, but it would be advisable to verify the BSB database, and make corrections for other mislabelling (easily identifiable by examining the database) before doing analyses."*

My response: The observation of a statistical balancing in the MicroPK database is due to a totally accidental statistical data formation. It may be observed as the consequence of the law of large numbers in large enough databases whose statistical average converges to null mean-shift. The statistical balancing of scores is neither a hypothesis under verification. It happened to appear in the large MicroPK database, in spite of the publication bias in it.

Before the publication of the MicroPK meta-analysis [Bösch et al 2006], there was a period of debate between a circle of researchers against its results (as in endnote #1) during which the reviewer, who seems to be well informed about the details of this meta-analysis, should have addressed such very serious concerns regarding the validity of data. If such objection had been raised, I understand that [Bösch et al 2006], who are all fluent in German and could easily spot such possible errors, would have duly corrected their database. As [Bösch et al 2006] have improved their meta-analysis following

the debates and have thus considered it correct for publication, then such belated critique addressed here is totally out of place.

16. *"I am puzzled by the expectancy MP proposal which I find odd for a number of reasons. First, if they were valid, PK-MP + experimental/control publication bias appear sufficient to reproduce the funnel plots in a qualitative sense. It seems ad hoc to add expectancy MP to this mechanism, the only motivation I can see being to lend an appearance of consistency to the RSA on the time-ordered BSB data, which I suggest the author has misinterpreted. Second, the research and publication process for studies is long and not sequential. Generally, the research for two separate and successive publications will have overlapped in time. How then does the expectancy apply? Third, since we know that publication bias is a prevalent and serious problem in many disciplines, should the "publication expectancy effect" only apply to PK studies? Wouldn't it also "statistically balance" studies of other psi effects, or any phenomenon with a small effect size? I find the ad hoc way in which it is used in the paper to be unconvincing. It is a central part of the paper, yet its basis in psychology is not reviewed, and the justification for applying the effect to the publication process is not developed at all."*

My response: The term 'expectancy MP' must refer to the 'experimenter expectancy effect' mentioned in my paper; the unintended and well-documented influence of the experimenters' hypotheses or expectations on the results of their research [Rosenthal 2004; Bakker et al, 2011]. My analysis shows that what the experimenters' report has been influenced by previous similar publications (Hurst exponent of the arranged MicroPK data being above 0.5).

- a. My paper does not propose, as the reviewer labels it, an 'expectancy Markovian process' (or, an 'expectancy MP').
- b. The 'experimenter expectancy effect' does not introduce a 'statistical balancing' of data, as it must be already clear by now.
- c. The 'experimenter expectancy effect' is not only present in the MicroPK studies, but is common in almost all areas of scientific enquiry [Rosenthal 2004; Bakker et al, 2011; R. Nuzzo, Nature, vol. 526, pp. 182-185, (2015)]. It is, therefore, not used 'ad hoc' in my paper.
- d. The 'experimenter expectancy effect' is very well reviewed within psychology, too [R. Nuzzo, Nature, vol. 526, pp. 182-185, (2015), see page 184].
- e. The suggestion that experimenters are in general influenced in reporting their results by previously published studies in the same field describes a trend. It does not imply that each and every experimenter has done so. That would be impossible anyway, for MicroPK studies that were published within the same issue of a journal or, in the same conference proceedings. That this tendency of experimenters is present in MicroPK tests is confirmed by the R/S analysis of the MicroPK time-series that yields a Hurst exponent above 0.5.

17. *"In addition, the author draws support for her arguments from a paper by Yu et al. that presents empirical reasons for rejecting the notion that consciousness is responsible for wave function collapse in quantum mechanics. One may accept that position without rejecting PK since we don't know if psi phenomena can be formulated within quantum theory or require an extension of it"*

My response:

(A). Regarding the sentence 'Accepting Yu et al. position without rejecting PK': The reviewer asserts that although consciousness may not be needed to collapse the wavefunction, it can perform the task, nevertheless. This assertion is fallacious as will be explained in (B) below. In any case, the MicroPK (and PK) hypothesis is not rejected by the paper of Yu et al., but by the strong evidence against it, as presented in my paper.

(B). Suggesting that consciousness can collapse the wavefunction, (that the mind can directly affect the physical process) is equivalent to suggesting that the MicroPK hypothesis is valid. But, there is no evidence in support of MicroPK. So, neither is consciousness needed to collapse the wavefunction, nor can it perform such feat.

(C). In order to formulate a (quantum) theory of a phenomenon there must pre-exist evidence to support it. As there is no evidence for MicroPK, there can be no theory (or, no extension of a theory).

Furthermore, the discussion is limited here to MicroPK alone and not to any ‘psi phenomena’, as the reviewer generalizes.

¹ Reexamining Psychokinesis: Comment on Bösch, Steinkamp and Boller (2006), Radin D., Nelson R., Dobyns Y., Houtkooper Y., *Psychological Bulletin* 2006, Vol. 132, No. 4, 529–532.

² As D. B. Wilson and W. R. Shadish admit in their commentary titled: *On Blowing Trumpets to the Tulips: To Prove or Not to Prove the Null Hypothesis—Comment on Bösch, Steinkamp, and Boller (2006)*, published in *Psychological Bulletin*, 2006, Vol. 132, No. 4, 524–528: “Bösch et al. did an admirable job searching for and retrieving all available psychokinesis studies, independent of publication status, and used well-justified eligibility criteria for establishing which studies to include in the synthesis”.