# MIND-MATTER ENTANGLEMENT CORRELATIONS:

# BLIND ANALYSIS OF A NEW CORRELATION MATRIX EXPERIMENT

# FINAL REPORT TO THE BIAL FOUNDATION FOR PROJECT GRANT 191/2016

Hartmut Grote[1]
[1]*Max-Planck-Institute for Gravitational Physics, Hannover, Germany*

## ABSTRACT

The work reported here is a rigorous conceptual replication of the so-called “Correlation-Matrix” experiment by an independent author, supported by Bial grant number 191-16. The experiment has been built from scratch with new hardware and software, testing 200 participants that have spent about half an hour each trying to ‘influence’ a physical random process visualized for feedback. The analysis software has been conceptualized following a strict blind analysis protocol. Blind analysis is a more rigid form of pre-registered analysis, in which the complete analysis software is written and tested before the data is actually analysed for the effect under study. The unblinding of the analysis, also called ‘opening of the box’ of the experiment described here has been performed live at the PA convention 2019 in Paris.
The main result was found to be not statistically significant and fell well within the expected random distribution of possible results. A second experiment (also following a blind analysis protocol) facilitating questionnaires and correlating them with the participants’ performance to ‘influence’ the physical random process (the main psi task) yielded a probability p=0.06 to have occurred by chance, under a null hypothesis. A post-hoc analysis of the hit rate for the psi task across all participants, which is mathematically independent from the correlation analysis, yielded a probability of p=0.06 as well, to have occurred by chance. Three unexpected anecdotal incidences that occurred during the execution of the experiment and the testing and actual analysis of the data may add to the canon of oddities and trickster-like effects sometimes reported in parapsychology research.

## INTRODUCTION

Mind-matter interaction or micro-psychokinesis (PK) experiments have a long tradition in Parapsychology and, due to inherent difficulties of these experiments, often have become more complex in their nature over the years. One of these more complex developments was to use *correlations* between psychological variables of a human agent with actual physical variables of a system. For example, instead of looking for an aggregate deviation from randomness one would correlate the output of a random number generator (RNG) with psychological traits such as belief in parapsychological phenomena, or others. An additional level of complexity was introduced by Walter von Lucadou by using *many* such correlations between different psychological and different physical variables within one experiment. Arranging all the resulting correlation factors in a matrix, he called this type of experiment “Correlation Matrix Method” (CMM) [11,12,14]. To evaluate statistical significance of a potential psi effect, all correlation factors in the matrix have to be evaluated together, as an ensemble.

The idea of the CMM method was motivated in conjunction with the hypothesis that psi may act in form of correlations, rather than in the form of a causal signal transmission. In an analogy to physics, such hypothetical entanglement-like correlations could not be used to reliably transmit information, which would make it unpredictable where in a given PK-experiment significant correlations would show up. Thus the combined result of many correlations is evaluated in a CMM-type experiment, without predicitng or expecting any particular correlation to show significance. In essence, this is a multiple analysis technique. The idea of a CMM-type experiment has also been related to the theoretical backgrounds of Generalized Quantum Theory [2] and the model of pragmatic information [13].

Von Lucadou performed three such CMM-type experiments and reported statistical significance on each of these [11, 12, 14], which more lately led to a new replication involving von Lucadou and other researchers [17]. There has been some debate about a statistical correct method to evaluate the significance of the ensemble of correlation factors in the matrix. At present, empirical methods (based on permutations and/or simulations of data) to estimate the statistical background distribution, as proposed and applied in [3, 4], seem the only viable method. The reason for this is that in virtually all correlation matrix experiments as performed to date, there are strong correlations between the psychological variables obtained, which renders all analytical statistical methods, which usually rely on statistical independence, invalid. The work in [17] has followed this empirical approach to estimate the statistical significance and obtained less significant results than were reported in [11,12,14].

In his two earlier experiments [11,12] von Lucadou used questionnaires to obtain psychological variables *before* the physical variables, derived from RNG output. In his later experiment [14], as well as in the replication [17], this design was changed to using the choice of button pushes of the participants as psychological variables. These variables from button pushes were obtained before, during, and after the generation of physical variables from RNG output. The latter has led to some controversy about possible causal correlations. While there seems to be no hint of causal correlations in the data of experiments described in [14] and [17], it seems more prudent to only use psychological variables that have been obtained *before* the physical variables, if truly non-causal correlations are to be elucidated and investigated.

In the study reported here, it was chosen to use the 'button-push' scheme, as in studies [14, 17], but to only use psychological variables obtained *before* physical ones for the main analysis. This is the main experiment of this study, as applied for with the Bial foundation and granted as 191/16. In addition to this main experiment, each participant was asked to fill a questionnaire before the apparatus-phase of the experiment. The questionnaires were planned to be correlated with the main RNG output across all participants, which constitutes the second experiment.

While the main experiment described here is a close conceptual replication of a CMM experiment facilitating the basic experimental design as used by von Lucadou and others, some notable differences will be pointed out. The second experiment is a conceptual replication of the work reported in [6].

## METHODS

### *Experimental Apparatus*

The experimental apparatus has been designed and built by the author of this study. It was chosen to build a new dedicated hardware device rather than using a standard computer, to give the experiment some feel of uniqueness. The device consisted of a case made of wood and metal, with a small (7 inch) high-resolution video screen and 4 dedicated push-buttons.

#### *Random event generator*

The apparatus featured a new hardware random number generator developed by the author. The source of randomness was based on thermal noise of two electrical resistors in a differential and shielded configuration that minimized coupling of environmental fluctuations to the noise currents of the resistors. Integer numbers of 16 bit lengths (random words) were obtained by 16 subsequent analog-to-digital conversions of the noise current with 10 bit resolution, using only bit 7 of each sample. To validate the quality of these random words, they were scrutinized with Marsaglia's "die-hard" suite of tests. No suspicious deviation from randomness was found.

To generate random events for the purpose of the experiment, a new random word would be drawn when required and compared to the previous drawn random word. If the new word was larger than the former, a logical '1' was defined as outcome. If the new word was smaller than the former, a '0' was defined. (The procedure would be repeated on a draw.) This procedure is similar to a 1-step Markov chain and is the procedure that was used for the experiments described in [14,17]. (The experiments in those references used single samples of noise of a Zener diode as inputs to the Markov chain.) We call the resulting bits 'Markov bits' or 'random events'.

#### *Two types of random events*

For about half of the participants of the experiment, the above procedure was used to generate Markov bits in real-time, when they were needed to progress the experiment. We call these bits RT-bits (Real Time). For the other half the above procedure was modified in two ways: First, the random words from the RNG were modified with an XOR operation with random (16 bit long) words from a pseudo-RNG algorithm (Mersenne Twister) which was seeded with independently drawn random words at the start of the experiment. Second, the full sequence of Markov bits was generated (using the XOR modified random words) at the start of the experiment and then stored in the memory of the control computer. Data was read from this sequence when they were needed to progress the experiment at later times, typically over a period of 15-30 minutes after their generation. We call bits generated in this way XPR-bits (XOR'd and Pre Recorded).

The split in these two types of random events was introduced to possibly test any result against these seemingly different sources of randomness. The analysis for the main and second experiment as defined in the blind analysis protocol did not make use of this distinction though.

## *Experimental sessions*

For each experimental session a participant was first instructed about the general goals of the study with the opportunity to ask questions. The study had been approved by the ethics committee of the Max Planck society in Germany, and each participant signed a consent form allowing the anonymous use of their data.

In a next step, each participant was asked to fill a 2-page questionnaire, composed of 50 questions to be answered on a 5-point likert scale. After this was completed, the participant was guided to the experimental apparatus located in a separate room on a desk in the given localities.

The participants were then asked to press a button to start the experiment. Upon this button press, the control computer would decide randomly (using the comparison of two random words) whether to use RT-bits or XPR-bits for this participant. The particpiants did not know of the two different bit generation modalities, and the experimenter (the author) did not know which bit type was determined for each participant until removal of the blind condition.

A test run was then performed in the presence of the author, with the author explaining the conductance of the experiment on the hardware device. After this procedure, each participant was left alone in the room to perform the main experiment in their own time.

A main experimental session consisted of 9 'runs' for each participant. For each run the participant was asked to use the left and right hand to push a button with either of the hand, respectively (there was a button for the left hand and a button for the right hand.) The possibility to push both buttons at the same time was left open and not deliberately mentioned. Each participant then had the task to press the buttons at a time of choice with each button push triggering the generation of an RT-bit at that time, or the reading of an XPR-bit from memory. The obtained bit determined the direction of motion of an arrow in the display (right or left). The participant had the task of trying to 'influence' the arrow to move more in the direction that it was pointing to, either the right or left hand side, or to keep it in the middle of the display, when it was pointing down. Each button-push was accompanied by a 'cheering' sound when the arrow moved in the intended (intention according to the instruction given) direction (a hit), and by less pleasant sound when moving in the opposite direction (a miss). When a new record position of the arrow to the right or left was achieved within a run (with a minimum of 3 steps to either side), a longer 'cheering' sound from a cheering group of people was played back, intended as a particular reward.

The condition of the arrow to point right, down, or left was given in each run, thus setting the nominal intention of the participant deliberately for each run. Each of the three conditions (right, middle, left) occurred 3 times, for a total of 9 runs. A run would be finished after 80 button presses and a new run would be started by the participant at will, upon the pressing of a third button.

After each run a short feedback was given in form of a text (as in [14]) and in form of the achieved hit scores (for the right and left conditions combined). After the 9 runs a summary feedback was given in form of a text as well as the total hit score and the information that the experiment had now finished.

*Data recording*

For each button press (80 per run) data was stored on the non-volatile memory of the control computer, namely:

- which button was pressed (left or right or both)
- time of button presses with a resolution of less than 1 microsecond
- raw 16 bit long word from 16 AD samples (for RT-bits) or 16-bit long word from memory (for XPR-bits)

In addition to this, three 16-bit long words, and correspondingly three derived bits, were stored as control data that were sampled appox. 1 ms before (1x) and 1 and 2 ms after (2x) the experimental sample as described above. This data was not planned to be used in any analysis, but to be used for consistency checks.

## *Participants*

Participants were recruited from different pools. A total of 200 participants, the pre-planned number, completed the experiment in the period from December 2017 to May 2019. Of these, 75 were recruited from the general population in the Santa Fe area in New Mexico, USA, and performed the experiment at Mountain Cloud Zen center in Sante Fe. 48 participants have been drawn from a student population of unspecified fields of study in Freiburg im Breisgau in Germany. These have performed the experiment at the IGPP institute in Freiburg. A further 56 participants have been recruited and tested in Berlin and Hannover in Germany, and the remaining 21 participants have been recruited and tested in Oxford and Cardiff in the UK.

The average participant age was 45.6 years with a standard deviation of 17.5 years. 78 participants reported as male and 122 as female. Figure 1 shows the age distribution of all participants.

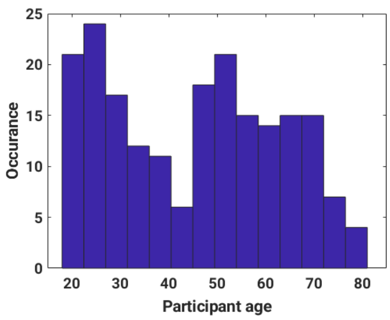


**Figure 1: Age distribution of all participants**

## *Blind Analysis Protocol*

Blind analysis is a strict form of a pre-defined analysis and is increasingly used in fields as medicine and physics [7] to prevent bias in the analysis procedure of a dataset. Less applied in social sciences and psychology, these fields may often benefit from blind analysis as well [8]. Potential bias in data-analysis of parapsychological experiments has been discussed in [1] and [15], among others.

In blind analysis, not only is the analysis plan pre-registered, but also the complete analysis algorithm, typically a computer program, is being written, fully debugged, and tested before the experimental data is analysed for the effect under study. This makes sure that detailed questions of the analysis that can affect the result (but that only may show up during performing the actual analysis) have to be decided beforehand, and any bugs associated with the analysis can be dealt with without introducing biases. In particular, during the 'blind' phase of the analysis development and testing, either artificial data is used, or the real experimental data is used, but in a disguised, i.e. blinded, form. The latter can conveniently be achieved in correlation-type experiments by permuting the association of psychological to physical variables between different participants. This was chosen for the development of the blind analysis code for this experiment.

The analysis for this experiment had been specified and tested during the data acquisition phase. In particular the data had not been looked at or analysed in any way before the unblinding, in agreement with the blind analysis protocol. The analysis of the experiment had been implemented in software (Matlab) as outlined below. The analysis code, as well as all the experimental data, had been pre-registered on the open science framework platform under the URL https://osf.io/xwhn6, but are embargoed until January 2021.

The data-analysis procedure follows closely the way the data was analysed for the latest experiments described in [14], except for the evaluation of statistical significance of the matrix of correlation factors. The choice of variables resembles the choice of the experiment reported in [17], with the exception that one of the psychological and physical variables each was replaced with new ones, as indicated below.

From the data recorded during the experiment (see above) the following 5 psychological and 5 physical variables were computed for each run:

Psychological variables:

- Number of left button presses
- Number of both button presses (simultaneous left and right)
- Number of times button (left or right) changed (This is a new variable not used before)
- Mean time between button presses
- Standard deviation of time between button presses

Physical variables:

- Number of hits (arrow moved in the intended direction). For the middle intention runs this was the negative number of steps away from the middle.
- Cumulative number of steps away from target
- auto-correlation
- number of cheer sounds (This is a new variable not used before)
- Standard deviation of the 80 raw random generator numbers

The 5 psychological and physical variables from the 9 runs yielded 45 psychological and 45 physical variables that were correlated against each other to yield 45 x 45 = 2025 correlation factors. Each correlation factor was calculated here across all participants. As mentioned above, for the main result of this study, only those correlations were considered that were based on time-forward correlations, i.e. only correlating psychological variables with physical variables from a later run. This resulted in 900 correlation factors remaining. While von Lucadou prefers to use all correlation factors, this author prefers the set reduced to time-forward correlations (i.e. psychological variables obtained before physical ones), to principally exclude the possibility of causal correlations.

After the matrix elements (the correlation factors) were obtained, a single test statistic was calculated for the matrix, combining all individual correlation factors. For the analysis described here, the absolute values of all correlation factors in the matrix were cubed and then the mean value of these cubed factors was computed. This way of analyzing the data is different from the method used in [11,12,14,17]: They used the number of 'significant' correlations in the matrix as the test statistic. That method has the potential disadvantage though that not all matrix elements can contribute to the result. Analysis of existing matrix experiments with both test statistics (performed by this author) show that the mean sum of absolute correlation factors cubed seems to be at least 'as good' as the alternative used by von Lucadou, i.e. yields results that are at least as significant. This was tested on the data of the experiment reported in [17]. A similar test statistic has been used in the work reported in [6], using the mean of squared correlation factors.

After determining the test statistic, in a second step the statistical significance of the obtained value was evaluated. This was done by empirically estimating the statistical background distribution using permutations of the psychological variables of participants against the physical variables of *other* participants. The test statistic for the un-permuted data was then compared to the distribution of test statistics form the permuted data. A one-sided hypothesis was used that the experimental data without permutations would show more and/or stronger correlations than most test statistics of the background distribution. The background distribution can also be simulated using artificial data as an alternative method. If both methods agree in their result, this can be taken as a strong argument for the validity of the estimated background, as pointed out in [4].

## Results

### *Main experiment*

The above pre-registered analysis was unblinded live during a presentation at the PA convention in Paris on July 6th, 2019.

Figure 1 shows the main result against the background distribution that has been obtained by multiple permutations of the experimental data set. The figure of merit falls well within the background distribution with p=0.76. The three control data sets obtained shortly before and after the experimental data also fall within the background distribution with p-values of p=0.74, p=0.22, and p=0.08, respectively (not shown in the graph).

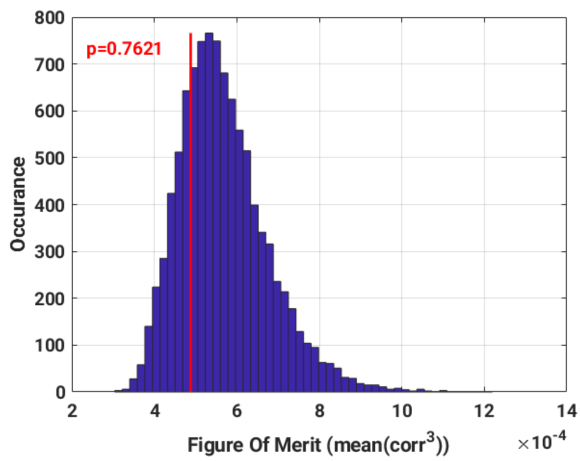


**Figure 1: Main result as obtained on the 'opening of the box event' during the PA convention 2019. The graph shows the statistical background distribution obtained with 10000 permutations of psychological against physical data (blue columns). The x-axis denotes the mean of all absolute values of cubed correlation factors in the matrix. The y-axis denotes the number of occurrences within the 10000 permutations. The red line shows the un-permuted experimental data. It falls well within the statistical value with a one-sided p-value of p=0.76.**

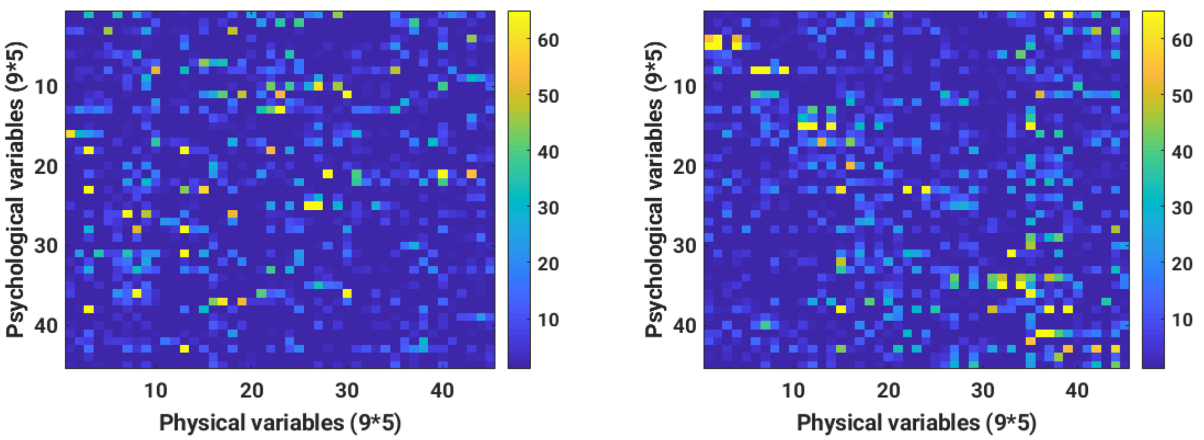


**Figure 2: Correlation matrix for a control condition (left) and for the experimental results (right). The color code denotes the cubed absolute correlation factors (times 10000 in units of the color code scales to the right of the graphs). Higher correlations can be observed close to the diagonal of the experimental matrix on the right. These are very likely causal correlations, which in any case had been excluded form the pre-planned main statistical analysis.**

Figure 2 shows the correlation matrices for one of the control data sets (left) and for the main experimental data (right). Notably there seem to be more correlations close to the diagonal in the experimental data matrix on the right. These are very likely due to causal correlations and are excluded from the main statistical analysis, which uses only the upper right part of the matrix.

However, when including the full matrix in the analysis, including the diagonal elements, the figure of merit of the experimental data results in a probability of p=0.17 to have occurred by chance under a null hypothesis. So even when including the potential causal correlations, the result is not significant.

*Second experiment*

The analysis for the second experiment was unblinded in the sole presence of the author on September 15th, 2019.

Figure 3 shows the result against the background distribution that has been obtained by multiple permutations of the experimental data set. The figure of merit falls slightly to the right hand side of the background distribution with a probability of p=0.064 to have occurred by chance under a null hypothesis. The three control data sets obtained shortly before and after the experimental data fall within the background distribution with p-values of p=0.61, p=0.53, and p=0.19, respectively (not shown in the graph). When splitting this analysis into the two types of random data used, the p-values are p=0.1448 for the real-time random data (RT-bits), and p=0.1678 for the XOR'd and pre-recorded random data (XPR-bits). This difference apparently is insignificant.

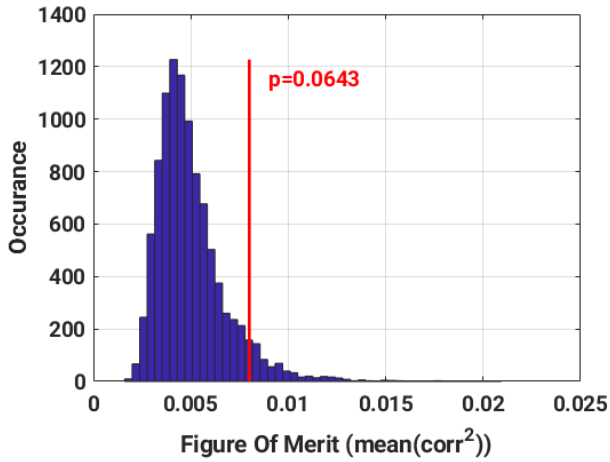


**Figure 3: Result of the second experiment as obtained on the 'opening of the box'. The graph shows the statistical background distribution obtained with 10000 permutations of psychological data of the questionnaires (50 variables) against physical data of the psi-task (1 variable). The x-axis denotes the mean of all (50) squared correlation factors. The y-axis denotes the number of occurrences within the 10000 permutations. The red line shows the un-permuted experimental data. It results in a one-sided p-value of p=0.064.**

*Post-hoc analyses*

Since one emphasis of this study was to use a blind analysis protocol, no particular post-hoc analysis was planned. However, blind analysis does not exclude post-hoc analysis in principle, which can range from consistency checks to exploratory analysis in search for novel effects in the data. Blind analysis just helps to draw a very clear line of distinction between pre-planned and post-hoc analyses.

Triggered by the first of the additional observation as reported below, it was decided to calculate the simple psi-taks related outcome of how successful the 200 participants were in 'shifting' the arrow in the display in the instructed left or right direction. It was found that the participants were 'successful' in the intended direction (hits) with a surplus of 138 bits out of N=96,000 trials (200x6x80). This corresponds to a probability of p=0.064 (one-sided) to have occurred by chance under a null hypothesis. Note that the standard deviation for the underlying Markov chain is SD=sqrt(N/12) and not the ususal SD=sqrt(4) as for independent random bits [11]. Also note that the numerical value for this probability (0.064) is the same as the above reported for the correlation analysis. This is pure coincidence, since the underlying degrees of freedom are completely independent in both cases (!).

The split analysis for the two types of random data yielded p=0.111 for the real-time data and p=0.107 for the XOR'd and pre-recorded data. Apparently both of these contributed about equally to the combined result.

Another post-hoc analysis was to run the analysis for the second experiment (the correlation between questionnaires and psi task) also with using the absolute value of correlation factor cubed, instead of squared. The result was exactly the same, i.e. a p-value of p=0.064. This nicely demonstrates the robustness of this figure of merit (at least for these data).

Further, the 50 psychological variables of the questionnaire were also correlated with the 5*9 physical variables as used for the analysis of the main experiment. This more resembles the way the data was analysed for the early CMM experiments, as of [11,12]. The result is not significant with a p-value of p=0.647.

## SOME ADDITIONAL OBSERVATIONS: TRICKSTER AT WORK?

I would like to report here three occurances that happened around the execution of the experiment and the testing and execution of the analysis, that I found quite remarkable. I would not classify these as 'post-hoc' analyses, but rather as observations that ocurred, without deliberatly looking for them.

*The first occurance* was the experimental result of an 8 year old boy, who performed the experiment as the only person besides the regular cohort of 200 participants. Since the experiment required the filling of a questionnaire and the signing of a consent form, it was only suitable for adults. However, during my presence at the house of his parents, the boy insisted of doing the experiment. So I let him have his way, deciding his data would not enter the analysis. At the end of his session, the boy had reached the most extreme result of the main psi task (the number of hits), compared to all 200 other participants. His total score was -19 with a two-sided probability of this or a more extreme result to have occurred by chance (under a null hypothesis) of p=0.0038. Obviously, if 200 participants do such an experiment, the occurance of one such result is not surprising at all. The surprise lies in the fact that this result was obtained by the only non-regular performer of the experiment.

*The second occurance* was related to the differential analysis, comparing the experimental results for the two cohorts of data using either the real-time RNG (RT-bits) data or the XOR'd and pre-recorded RNG data (XPR-bits). This analysis was tested using artificial pseudo-random data generated with Matlab. The first three tests of this analysis code yielded p=0.06, p=0.95, and p=0.0003 for the results (or more extreme ones) to have occurred by chance. As a sort of memo, Figure 5 shows the last result as it was first obtained. The 'p=0' result of this test came from the limited number of permutations of n=1000 used for estimating the statistical background of the test data. In a subsequent run using n=10000 permutations for the same test data, the more accurate estimate of p=0.0003 as reported above was obtained. The combined p-value for the three extreme results is of order p~0.0001. Further tests with new artificial data converged to a uniform distribution of p-values as expected. So while this was a test run for the analysis code, subsequently nothing was found wrong with the code.

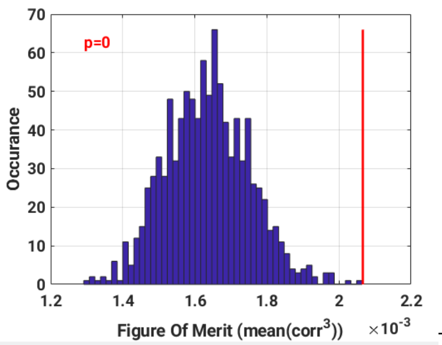


***Figure 5: Result of third test run for the split RNG analysis. The author was struck by the extreme result, with subsequent results converging to the expected uniform distribution.***

*The third occurance* was related to finally performing the differential analysis on the unpermuted data, once the testing of the code for the split RNG analysis had completed. Figure 6 shows the two histograms of background distribution and the results for the two random event groups. After I had first produced the graph on the left hand side I thought that the composition of digits of the obtained p-value of p=0.5876 looked 'funny', like a sequence of numbers after the decimal point, though shifted by one digit. While briefly wondering about this, I produced the graph on the right hand side next, which, to my surprise, resulted in a p-value of p=0.4321. Note that these 2 p-values are completely independent from each other, since they represent the data of the main experiment, split into two independent groups, according to which type of random events was used.

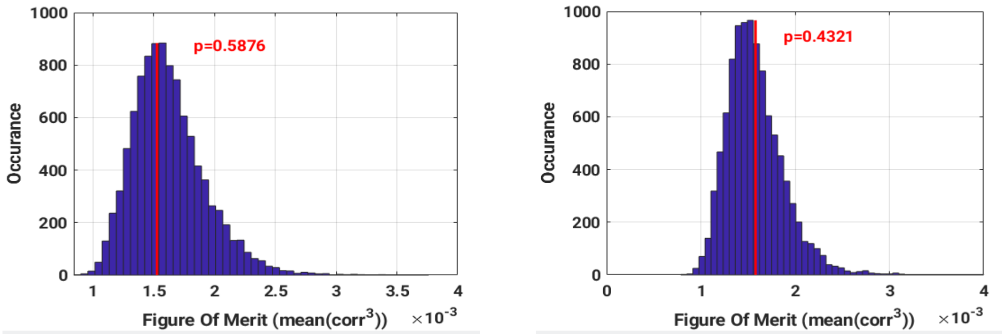


**Figure 6: Result of the main experiment as split for the two types of RNG data used: Using the real-time data (left) and the XOR'd and pre-recorded data (right). The author was struck be the 'funny' numbers of the estimated probabilities.**

## DISCUSSION

Each reader may draw their own conclusions, but the experiment and its results reported here may merit discussion of several topics: 1) The conceptual replication of a CMM experiment using a blind analysis protocol. 2) The conceptual replication of a CMM-similar experiment correlating questionnaires with a single psi task, also using a blind analysis protocol. 3) A post-hoc finding of marginal significance of the main psi task and its relation to the replication problem and the idea of the CMM experiment. 4) Trickster-like properties of the experimental outcomes and incidental observations. And 5) some thoughts about future research.

1) The conceptual replication of the CMM-type experiment of the type performed before in [14] and [17] (main experiment as reported above) has yielded no evidence of anomalous correlations. While the methodology of using blind analysis was slightly stricter than the methodology in [14] and [17], I do not believe that this was a decisive factor in this outcome. If we do not question the positive results reported in [11,12,14] and [17], I would interpret the null result of the experiment reported in this study as either the result of a decline of effects in repeated parapsychology experiments, or as a sign that I am not 'gifted' as psi-conductive experimenter (at least for this experiment). Both of these interpretations touch upon profound problems of parapsychological research.

If experimenter-psi effects are taken into account (see [10] for a more recent overview), then in addition one may also consider the audience that observed the unblinding of the main experiment during the presentation at the PA convention in Paris. With a notion of humor I would say I am a bit disappointed by the psi performance of this particular, select, audience in 'influencing' the outcome of the main experiment. In a snap poll about half of the audience expressed the opinion that the result of the study was not yet determined at the moment before the unblinding of the analysis, leaving room for the audiences to 'influence' the imminent outcome. In their defense one may argue that the total person-hours of the audience are significantly less (~100) than the person hours the author invested into this study (~600).

While the experimenter psi hypothesis is often neglected in parapsychological research, perhaps to no small degree because it renders research even more difficult in terms of reproducibility and other factors, a correlation experiment seems to be particularly prone to experimenter-psi, if it exists. This is because in a correlation experiment, no single individual participant can be 'successful' on his or her own. The result of the experiment is always one (or many) correlation factor(s) obtained across an *ensemble* of participants. For a single participant the result would simply not be defined. The group of participants can generate the effect under study only as an ensemble. So in this context its completely unclear what the role of a 'gifted subject' may be. One can speculate that if experimenter-psi exists (and why would it not under a psi hypothesis?) then a correlation-type experiment is particularly prone to experimenter-psi.

2) The second experiment in this study, a conceptual replication of the work reported in [6], yielded marginal evidence (p=0.064) of anomalous connectivity – of the same order of magnitude as found in [6]. It is worth noting that this way of performing a CMM-like experiment is closer to the early experiments of von Lucadou as reported in [11,12] where also questionnaires were used. I would not claim here that a new experimental paradigm has been found, but rather take note that this is only the first replication of the work in [6], and effects may decay in the future. Taking the marginal evidence serious, another interpretation would be that of a Trickster-like result, where the effect went back to where it was found in the past, i.e. in the correlation studies using questionnaires as reported in [11,12]. Note though that only one physical variable was used in the planned analysis presented here.

3) The post-hoc result of p=0.064 (same number as above, but different result!) for the main psi task of 'influencing' the left-right direction of the observed arrow by the participants may be a chance result, of course. It is remarkable though that such an effect shows up in a place where it was not expected, since the whole CMM idea was developed around the notion that psi would not act in a 'signal-like' fashion such as to 'influence' the outcome of RNG's in a particular direction. If one interprets the observed p=0.064 deviation as hint of an anomaly, then the whole history has come full circle: An effect in a more simple RNG-pk 'influence' type of experiment was observed, but not in the more complex CMM-type analysis. Again, this has more of a Trickster quality than anything else. The CMM method was an ingenious idea to reliably observe anomalous connections in data by way of multiple correlations. However, we may be facing Trickster qualities that will always be a bit *more* ingenious. A conclusion that certainly would concur with J.P. Hansens analysis of parapsychological research [5].

4) Further to the Trickster-like observation in 3), the three reported incidental observations can be seen as pure chance, and would probably be dismissed and not be reported by most researchers, certainly so in mainstream academia. In the context of a psi experiment, to me they had a very special quality, and I am more inclined to interpret them as another Trickster manifestation. While psi effects have been sought for in a carefully planned and executed experiment, they have been found at the edges of this endeavour, in unexpected places.

5) My main conclusion would be that the CMM idea will not escape the replication problem of parapsychology. Another possible interpretation would come from dominant psi-experimenter effects. Under a psi hypothesis, psi-experimenter effects need to be taken serious. This has been pointed out by many, see [10] for an overview, but still is not a theme frequently adhered to across the board. In addition to psi-experimenter effects, the Trickster theme (which obviously would also play into psi-experimenter effects) may be more ubiquitous than existing scientific literature may make us believe.

The experiment reported here has spanned methodology from the extreme controlled (blind analysis) to the anecdotal (subjectively 'funny' or surprising incidences) and I believe the full breadth of these methodologies is needed to shed light on psi in the experimental domain. Obviously it would make not much sense now to propose an experiment where we look for a Trickster effect, such as for odd statistical results when testing analysis code. This would lead to a regression ad infinitum. In a sense this is what the CMM idea was trying. We may as well continue to do so, but I believe it will be important to pay attention to the more subjective qualities of such experiments as well. I strongly seems to me that experimental parapsychology research should be rigorous (to not fool ourselves) and descriptive (in a subjective experience sense) at the same time. A planned experiment then resembles also something of a single case study in the field. This avenue has been less explored, but the recent work by Herb Mertz is a good example in this category [9].

## References


1. Bierman, D. J., Spottiswoode, J. P., & Bijl, A. (2016). Testing for questionable research practices in a meta-analysis: An example from experimental parapsychology. *PLOS One*, 11(5):e0153049.
2. Filk, T., & Römer, H. (2011). Generalized quantum theory: Overview and latest developments. *Axiomathes*, 21: 211–230. arXiv:1202.1659
3. Grote, H. (2015). A correlation study between human intention and the output of a binary random event generator. *Journal of Scientific Exploration*, Vol.29, No.2, p.265-290.
4. Grote, H. (2017). Multiple-analysis correlation study between human psychological variables and binary random events. *Journal of Scientific Exploration*, Vol.31, No.2, p.231-254.
5. Hansen, G.P. (2001). The Trickster and the Paranormal. *Xlibris,* ISBN 1-4010-0082-7
6. Jolij, J. and Bierman, D. (2019) . Two Attempted Retro-Priming Replications Show Theory-Relevant Anomalous Connectivity. *Journal of Scientific Exploration*, Vol.33, No.1, p.43-60.
7. Klein, J.R. and Roodman, A. (2005). Blind analysis in nuclear and particle physics. *Annual Review of Nuclear and Particle Science*, Vol 55: 141-163.
8. MacCoun, R. and Perlmutter, S. (2015). Blind analysis: Hide results to seek the truth, *Nature*, Vol. 526, Iss. 7572.
9. Mertz, H. (2020). The selection effect: How consciousness shapes reality. *Penn Wolcott Press,* ISBN 1733508007
10. Palmer, J. and Millar, B. (2015). Experimenter effects in Parapsychology research*, in: Parapschology A handbook for the 21st century (Eds. Cardena, E., Palmer, J., Marcusson-Clavertz, D.), Mc Farland & Company*, p. 293.
11. Von Lucadou, W. (1986), Experimentelle Untersuchungen zur Beeinflussbarkeit von stochastischen quantenphysikalischen Systemen durch den Beobachter, *H.-A. Herchen Verlag: Frankfurt*.
12. Von Lucadou, W. (1991). Locating Psi-bursts – correlations between psychological characteristics of observers and observed quantum physical fluctuations”, *in: Delanoy, D.L. (ed.)* The Parapsychological Association 34th Annual Convention, Proceedings of Presented Papers. Heidelberg: Parapsychological Association. pp. 265-281.
13. Von Lucadou, W. (1995). The Model of Pragmatic Information (MPI), *European Journal of Parapsychology* 11, 58-75.
14. Von Lucadou, W. (2006). Self-Organization of Temporal Structures – A Possible Solution for the Intervention Problem. *Daniel P. Sheehan (ed.): Frontiers of Time. Retrocausation - Experiment and Theory. AIP Conference Proceedings*, Vol. 863, AIP, Melville, New York, 293-315
15. Wagenmakers, E. J., Wetzels, R., Borsboom, D., Kievit, R., & van der Maas, H. L. J. (2015). A skeptical eye on psi. In *Extrasensory Perception: Support, Skepticism, and Science edited by E. May and S. Marwaha*, New York: Praeger, pp. 153–176.
16. Walach, H., von Lucadou, W., & Römer, H. (2014). Parapsychological phenomena as examples of generalized nonlocal correlations: A theoretical framework. *Journal of Scientific Exploration*, 28(4):605–631.
17. Walach, H., Horan, M., Hinterberger, T., & von Lucadou, W. (2020). Evidence for anomalistic correlations between human behavior and a random event generator: Result of an independent replication of a micro-PK experiment. *Psychology of Consciousness: Theory, Research, and Practice*, Vol.7, No.2, pp. 173-188