Journal of Scientific Exploration, Vol. 35, No. 4, pp. 788–828, 2021 0892-3310/21

RESEARCH ARTICLE

Nailing Jelly: The Replication Problem Seems to Be Unsurmountable—Two Failed Replications of the Matrix Experiment

Harald Walach

Poznan Medical University, Department of Pediatric Gastroenterology, Poznan, Poland;
University Witten/Herdecke, Department of Psychology, Witten, Germany;
change Health Science Institute, Berlin, Germany
hwalac@gmail.com, https://harald-walach.de +49-30-46906958

Karolina A. Kirmse

Technical University Chemnitz, Department of Psychology, Chemnitz, Germany

Peter Sedlmeier

Technical University Chemnitz, Department of Psychology, Chemnitz, Germany

Hans Vogt

Change Health Science Institute, Berlin, Germany

Thilo Hinterberger

Medical University Regensburg, Department of Psychosomatic Medicine, Regensburg, Germany

Walter von Lucadou

Parapsychology Counselling, Freiburg, Germany

Submitted December 17, 2020; Accepted September 12, 2021; Published December 30, 2021

https://doi.org/10.31275/20212031

Abstract-—We have reported previously on positive effects found in the matrix experiment (Walach et al., 2020). This is a setup where a random event generator (REG) drives a display, which participants are instructed to “influence” at will, i.e., in a psychokinesis (PK) setup. The difference of this matrix experiment from standard micro-PK REG experiments was that instead of the deviation from randomness, a large array of 2025 correlations between the behavior of the participant and the behavior of the REG was tested. This previous experiment was significant, and we devised a consensus protocol, which was deposited before commencement, according to which we conducted two independent replications with the same experimental setup and equipment. In the first experiment 64 participants conducted the experiment in one location under the experimental guidance of KK (power = 0.88), in the second experiment 40 participants conducted the experiment in another location under the experimental guidance of HV (power = 0.69). The analysis used a nonparametric randomization test with 10,000 iterations. Neither of the two experiments was significant. While in the first experiment a very small, but non-significant effect was found, in the second experiment no effect was detectable. We discuss the findings in the context of the larger debate around replicability of parapsychological (PSI) research results and our theoretical model. The replication problem and this failed replication is likely part of the systematic nature of such effects. This makes it unlikely that experimental research alone will be successful in the long run in demonstrating PSI effects. Our conclusion is that the matrix experiment in and of itself is not a replicable paradigm in PSI research.

INTRODUCTION

Parapsychologists were probably among the first social science researchers to understand the necessity to publish all negative findings and to insist on replications. Replications come in various forms: Identical replications are rather rare and refer to the replication of the very same experiment, including all materials, methods, procedures, and statistical analyses except that either the same research group or other researchers run a set of different subjects. Mostly, however, replications also vary some kind of procedure, thereby making replications conceptual (Schmidt, 2016). We established previously an experimental paradigm which we hoped would lend itself to replication (Walach et al., 2020). We report here on two replication experiments that failed to replicate the previous result.

Parapsychology and Replication

The idea of replicating experiments is, of course, to exclude accidental findings, false positives, and small statistical fluctuations being identified as systematic effects. Physicists have long been taken as examples of doing natural science rigorously, and they demand at least 5 sigma or standard deviations of the standard normal curve and associated probabilities as confirmations from various experiments until they accept an effect as veridical (Abbot & LIGO Scientific Collaboration and Virgo Collaboration, 2016; Grote, 2018; Horton, 2015). They usually achieve this by accumulating data through multiple observations or through multi-lab experiments, as in the LIGO experiments that detected gravitation waves. Thereby, the very same experimental setup, including all analysis pipelines and data-collection procedures are standardized and logged. The idea behind this procedure is that experiments are detectors for stable and local signals that may be weak but can be eventually separated from noise. Local signals travel at, or slower than, the speed of light and appear regular, i.e., exhibit lawful behavior. The laws might either already be known and predict certain signals, as in the case of gravitation waves that were long predicted by the standard model of cosmology, but the signals might be very weak or very rare, like gravitation waves, or researchers might surmise that unknown laws underly hitherto undetected local signals. This is what some researchers in the parapsychology research community assume (Carr, 2015a; May et al., 1996; May et al., 2018; Radin, 2018).

In parapsychology we are looking for effects whose nature we do not understand, because there is no accepted theory in the first place. Some theories with testable consequences have been proposed, such as “decision augmentation theory” (DAT), which supposes that all PSI effects are precognitive effects, where anomalous cognition of future events is used to augment decision making (May et al., 1995, 2000, 1996). Apart from the fact that this model has problems explaining makro-PK, spuk, and poltergeist phenomena, it also has some empirical evidence against it (Dobyns & Nelson, 1998). Observational theories use some form of argument from a von-Neumann–Wigner-type of interpretation of quantum physics, in which human consciousness is central in collapsing the wave function (Houtkooper, 2002; Walker, 1975, 1979, 2011 [1974]). In this family of theories, it is the joint observational effect of those looking at the data that actually produce the effect. While this might be a theoretical option, it depends on the interpretation of the measurement process and the acceptance of a dualistic model, none of which are currently universally accepted

(Bierman, 2010). Finally, a model very similar to ours is the Conscious Induced Restoration of Time Symmetry (CIRTS) model (Bierman, 2010). This starts from the assumption that most physical theories, with the exception of thermodynamics and special relativity, are time symmetric, and that the brain sustaining consciousness might be a system that restores time symmetry through providing time-negative effects as in precognition. The CIRTS model assumes that there is a kind of signal or informational element in PSI-effects that are, however, bounded by physical theory. The model which we favor (see below) is also derived from physical theory and hence a potential scientific candidate for PSI-modeling, but more strongly than all other models assumes that signal coding is strictly prohibited.

Are PSI effects due to causal and local signals, i.e., obeying the special theory of relativity? Are the signals of a known type, i.e., belonging to the four types of exchange particles of the four basic known forces in the universe, for instance are we looking for photons as exchange particles of the electromagnetic force, or for other particles (Penrose, 2004), or are we looking for completely different, yet nevertheless local causes and signals of a physical nature? Or are we even looking for completely different types of signals that cannot be encompassed within the standard model of physics and hence would, if discovered and proven as stable and replicable, entail a widening of our worldview similar to that produced by the advent of quantum mechanics? Some physical concepts that use higher dimensional models of space and time than relativity theory and quantum theory would suggest this (Carr, 2015a, 2015b; Heim, 1989).

The parapsychological database is jagged so far. While we do have many extremely intriguing phenomena on a phenomenal level (Braude, 1986, 2017; Grosso, 2016), strong and well-documented cases, and highly significant meta-analyses summarizing research fields or experimental paradigms across researchers, variations, and time (Cardeña, 2018; May & Marwaha, 2018, 2019a, 2019b), critics are also correct in pointing out that it is not possible to name one single parapsychological experiment as foolproof and resistant to experimental replication (Alcock, 2003; Reber & Alcock, 2020). It is also true that we have a replication crisis in psychology in general, i.e., the inability to externally replicate experiments that were thought to be proven (Open

Science Collaboration, 2015; Schooler, 2011). This replication crisis affects all sciences, according to a survey where 90% of polled scientists say that there is a problem with replicability (Munafò et al., 2017). It is notorious in medicine despite the fact that medical interventions are widely used and believed in (Horton, 2015; Ioannidis, 2005). So why bother about the lack of replicability in parapsychology? Perhaps parapsychology is even more replicable than standard science, but only more controversial and hence less accepted (Radin, 2018; Schwartz et al., 2018)?

Assumptions Behind Replicability, Synchronicity, Generalized Quantum Theory, and Generalized Entanglement

Replicability, we observed above, makes the implicit assumption that we are dealing a) with local, causal signals that are b) regular, following some lawful rule and c) are therefore always available for experimental control and manipulation. In final consequence they would be amenable to human engineering, once the rules and the lawful behavior are fully discovered. We started our research from the assumption that the *lack* of replicability is part of the *systematic* nature of parapsychological effects. In other words, we assumed that the effects of parapsychology might be *lawful*, *but not* of a *local-causal* nature. This sounds like a contradiction, but it is not. The focus of science has so far been mainly on local-causal signals, because once they are discovered, they can be put to use: We have used electricity, once we discovered the nature of the electromagnetic force. We can use the gravitational force, e.g., by sending satellites into orbit. We even made use of the strong atomic force that keeps atoms together, when we started to engineer atomic fissure. We make use of the knowledge of the weak force in isotope calculations, Geiger counters, and the like. So, what would be a lawful, yet not causal-local event?

In the realm of psychology, Carl-Gustav Jung and Wolfgang Pauli, the physicist, discussed exactly such phenomena, i.e., lawful yet not causal relationships, under the umbrella term "synchronicity" (Atmanspacher & Primas, 2006; Atmanspacher et al., 1995a; Jung, 1952; Mansfield, 1995; Peat, 1992). These are material events in the material world that in their occurrence appear to be without cause, i.e., they happen "accidentally" or "randomly", yet they have a correlation with the psychological state of a person who relates to those events. Someone calling on the phone, while a person is desperately in need of this contact would be a typical example. This phenomenon has also given rise to a series of experiments (Sheldrake & Smart, 2003) which, in our view, demonstrate that this phenomenon actually exists but is not of a causal nature (Schmidt et al., 2004a), exactly as Jung and Pauli would have postulated.

Jung and Pauli wanted synchronicity to be seen as a type of lawful relationship that is complementary to causal relationships. They codified this in their famous quaternity (Jung in a letter to Pauli on November 30, 1950, in Meier, 1992, p. 64; Meier, 2001), depicted in Figure 1. This also implies that synchronistic relationships that are due to psychic meaning-making or constellation of an archetype, as Jung called it, and are in a way part of a "deeper" structure of reality than local causes. For synchronistic, correlational relationships are part of this primordial level of "indestructible" energy. Similar to this ontic level of "indestructible energy", that some physicists call the endophysical level of unbroken unity (Atmanspacher et al., 1995b; Primas, 1994a, 1994b), there are also relationships that pertain to this level and might be put to use (Lucadou, 2019). They might be lawful, but they are not causal in nature. The causality principle only operates on the level of the space–time continuum, or on the level of exo-physics, where clear delineations and determinations can be made, because the original unity is broken into measured and measurable parts.

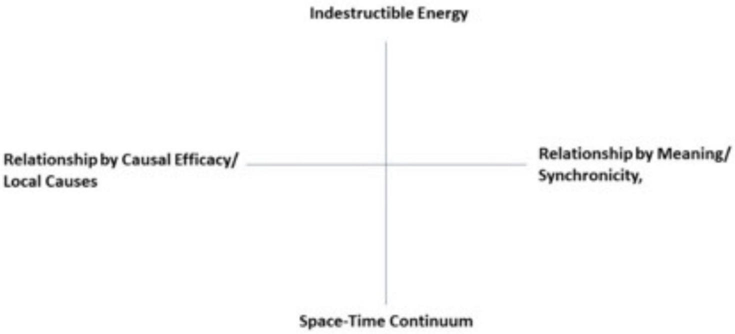


Figure 1. **The quaternity Jung suggested to Pauli: Local causes are complementary to correlational relationships or synchronicity in a similar way as the space-time continuum is complementary to indestructible energy.**

In that sense, attempts to uncover a purported causality in this realm is futile, simply because there are no local causes operative here, but only formal and final causes, to speak in Aristotelian terminology. Another way of putting this is that there are likely only correlations of a lawful but not causal nature. Indeed, there are examples in the physical world for such a type of relationship as well, namely quantum entanglement correlations (Atmanspacher et al., 2002; d'Espagnat, 1997; Schrödinger, 1935; Shimony, 1989; Stillfried, 2010). These are quite lawful, but not causal in the sense that the lawfulness of these correlations is not mediated by any exchange particles of force or energy. This is something that Einstein had already observed and this is the reason why he called them "spooky actions at a distance" (Einstein et al., 1935). They remained purely hypothetical for a long time, derivable from the formalism, but no one knew whether they are "real". This dispute was settled after John Bell derived his famous inequalities as a boundary condition for joint probabilities that are mutually exclusive, obeying locality conditions (Bell, 1987). This inequality gave rise to an operationalization and an experimental test which finally clarified the issue (Aspect et al., 1982a; Aspect et al, 1982b). There are indeed non-local correlations, i.e., lawful, yet not causally mediated regularities in nature.

Physical quantum correlations have been empirically documented as factual beyond any reasonable doubt (Handsteiner et al., 2017; Ma et al., 2012; Salart et al., 2008; Stefanov et al., 2002). Meanwhile, they are the basis for various new fields of research and application, from quantum computing to quantum encryption. And some think they are the basis for our neural operations as well (Hameroff & Penrose, 2014). Therefore, it might be reasonable to assume that such lawful, yet a-causal and non-local relationships could also play a role in the wider area of human affairs or in macroscopic nature. However, this would necessitate that either physical entanglement correlations that are normally only detectable under highly controlled and artificial conditions can also be preserved to some degree in the macroscopic environment; or that there is an equivalent to physical entanglement correlations that are exactly those meaningful correlations Jung spoke of, but not necessarily of a physical nature. Such correlations might be, for instance, systemic, i.e., pertaining to the general setup of a system of different physical constituents, and not only strictly physical in nature (Atmanspacher et al., 2002; Lucadou, 1995, 2015b). This is the path some of us have chosen, in assuming that there is a generalized form of entanglement that is operative in various types of systems, provided they have a certain structure (Filk & Römer, 2011; Walach & von Stillfried, 2011a, 2011b). We assume that parapsychological effects are due to such correlations, lawful, yet not causal, regular, yet not local (Lucadou, 2015b; Walach et al., 2014).

The No-Signal-Transfer (NT) Axiom and the Development of the Matrix Experiment

A corollary of this assumption is that if such correlations are mistaken for causal-local regularities and could be potentially used as such they will either change channel, i.e., show up in the control condition, or they will reverse signs, i.e., become significantly weaker, or are seen in different parameters. The reason for these observations is given by the fact that physical entanglement correlations must not be used as causal signals, and this can be formally proven (Lucadou et al., 2007). We therefore *assume* that this no-signal-transfer axiom (NT-axiom) also holds in the generalized case, although here it cannot be proven to be true, but is assumed to hold.

This NT-axiom states:

> *If a system is governed by non-local correlations but is treated as if the correlations were local causes, and if a signal is extracted from it, or could be extracted in principle, then those purported signals will break down in a second experiment, or when so used.*

This means that experiments on systems constituted by such non-local correlations that are repeated, constitute a violation of the NT-axiom and are likely to demonstrate a breakdown of such effects, either by a dwindling of the effect size or by demonstrating paradoxical effects, such as having the effect show up in the control group, or changing its sign. It also means: This is only applicable to *replications* of experiments. We will return to this problem in the Discussion.

Indeed, there is a series of empirical hints that testify to the ubiquity of this phenomenon in parapsychology. A recent example is a commissioned identical replication of a previously reported experiment in which mental effort of trained meditators was supposed to affect an interference pattern in a standard double-slit optical setup (Radin et al., 2012). The strictly preregistered and controlled study came out negative (Walleczek & von Stillfried, 2019), although an effect can be seen in a completely different channel, in the variance. The same was found in the multisite replication of the PEAR Lab's micro PK-experiment (Jahn & Dunne, 1987). Although a case can be made that the PEAR Lab's database was largely due to the effect of some gifted subjects, the consortium replication between Princeton, Freiburg, and Giessen was predefined as a large replication study of the PEAR Lab procedure in a protocol, in which Walter von Lucadou also predicted the negative result.

The replication was negative (Jahn et al., 2000), but secondary parameters that were not logged in the protocol, variance and nonlinearity parameters, were clearly significant (Pallikari, 2001). Maier and colleagues had the same experience in a series of PK- and priming experiments (Dechamps & Maier, 2020; Maier et al., 2014; Maier & Dechamps, 2018; Maier et al., 2018). This has also been observed in other datasets from parapsychology (Bierman, 2000).

We started from the assumption that, if this lack of causal stability is to be expected there might be a workaround by testing for some kind of indirect parameter that would prohibit the coding of a causal signal. Standard experiments that use a control group are, by default, cause detectors and thereby allow the coding of a signal. For instance, they yield a result which in a strict replication experiment could be used to code a signal: Scores above or below the mean of the first experiment would be the signal to code a 1 or a 0. This, however, would violate the NT-axiom and hence would cause the effect to vanish or change track, if our model is correct. Therefore, we sought an experimental setup that would be as immune as possible to such potential violation.

One such setup was developed by Walter von Lucadou in what he called the "Matrix Experiment", meanwhile referred to as the "Correlation Matrix Method – CMM" (Lucadou, 1974, 1986, 1987a, 1987b, 1991, 2006, 2015a; Lucadou et al., 1987). The idea here is not to define a clear outcome parameter which would be prone to violating the NT-axiom, as it would allow for signal coding, but to use an array of variables in a correlation matrix. The correlation matrix reflects the correlation of the interaction of human intentions or human behavior with a physical system that is otherwise locally decoupled. In our case the physical system was a micro-psychokinesis (micro-PK) experiment. A computer displayed a fractal, a Julia-set, whose change—growth or shrinkage—was driven by a random event generator (REG). That means the behavior of the fractal could not have been influenced by ordinary means of interaction. However, human participants were instructed to do so by their intentionality and were asked to move the sampling process of the random event generator forward by pressing either one of two keys on the keyboard of the laptop computer that ran the experiment. These keystrokes represent the psychological or behavioral variables, while the behavior of the physical system represents the physical variables. These variables can be correlated across all experiments and all participants and yield a correlation matrix. If the correlation matrix contains a signature of the intentional effect or the entanglement effect of participants with the experiment or physical system, then we would expect more significant correlations than by statistical chance expectation or in a control experiment that is run without a participant present.

Indeed, von Lucadou's previous experiments were supportive of this idea and produced more significant correlations than expected by chance and more than seen in a control matrix. Thus, we set out to replicate this experimental setup with a larger, well-controlled experiment. We rebuilt the hardware and software—the random event generator and the control software—from scratch and enlarged the matrix into a matrix of 45 psychological and 45 physical variables (because there were 5 such variables per run and 9 runs made up an experiment), yielding a matrix of 2025 cells. We created a robust non-parametric system of statistical evaluation by simulating 10,000 such experiments and deriving the statistical significance from it. This first large replication in two labs yielded a significant but fragile effect, as significance broke down in reasonably improved methods of analysis (Walach et al., 2020).

We then convened an international consortium of experts to arrive at a consensus protocol. This protocol followed our original one quite closely with a few exceptions (see Methods section below) and formed the basis for future replications. One such replication was conducted by Karolina Kirmse as part of her master's thesis under the supervision of Peter Sedlmeier. Another replication was conducted by Hans Vogt and Harald Walach. Both replications came out negative. We report on these replications in this paper and will end with a few ideas about potential ways forward and why we think it will be a difficult challenge to experimentally prove anomalistic effects using experimental models (Rabeyron, 2020).

METHODS

We used a predefined protocol that was the result of a consensus meeting of experts. The studies reported here are in fact replications of the parent study (Walach et al., 2020). The protocol was defined and published beforehand on the Open Science Framework platform (https://osf.io/cx2tf/). Since it is described there in detail, we will only summarize the most important elements here. The experiment is a comparatively strict replication of the parent study, as the same equipment, the same material, and the same procedures were used with only a few exceptions that are described below. The criterion for a successful replication was a significant result as determined by a statistical randomization test (see below). Since we do not assume a stable, causal effect, a standard power analysis is not part of the protocol but can only be provided as a post-hoc analysis.

Material and Participants

We used the same equipment as in the parent experiment. KK was lent one of the four REGs that were used for the first experiment and received a copy of the software program that operated the experiment. This software program was custom written in C following the first code which was programmed in Basic. This program operated the experiment automatically, prompted the experimenter and the participants for inputs, and wrote the data into a file.

The first replication was conducted by KK in Dresden with a broad group of participants recruited mainly in public spaces, the second by HV in Witten with a group of students of psychology, gaining course credits through their participation. The experiment was advertised as an experiment in extraordinary facilities and was conducted face to face, one after the other. Before the experiment started, the experimenter switched on the computer and the equipment with a lead time of half an hour to allow for drift and warming up.

The experimenter greeted the participant and briefly explained the experiment and handed out a consent form, as well as a very short questionnaire. The questionnaire data were deliberately not used, as in the previous experiment, but had the function of involving the participant with the experiment. In Experiment 1, however, this questionnaire was extended and evaluated for exploratory purposes (see below). When the participants were ready, the experimenter started the program and left them alone. The participant could take as long as necessary. They had the instruction to "influence the movement of the fractal on the screen" in the indicated direction and knew that they had to press either of two shift keys on the computer keyboard to move the sampling process forward. Each time either one of these shift keys was pressed, the REG was sampled and the result was used to generate a movement of the fractal displayed on the screen. The sampling process was filtered by a Markov chain instead of the frequently used XOR-filter. This was done for two reasons. First, Markov-chain filtering makes a process smoother and look more natural. Most natural processes, like the weather, are Markov processes, i.e., they contain one or two lags of memory. Second. the Markov process preserves some of the physical properties of the REG. Perceptually, this resulted in the appearance of a very smooth movement of the fractal. What the participant did not know was that when both shift keys were pressed, the sampling process would go on until one of them or both keys were released. The sampling process was repeated 80 times, since each sub-run consisted of 80 such trials, and 3 sub-runs with three different instructions made up one experiment. Each run was associated with a specific instruction to either grow or shrink the movement of the fractal or keep it constant. These instructions were conveyed by red arrows on the screen, and each instruction was repeated 3 times at random. Thus, a full experiment consisted of 3*3 sub-runs with 80 trials or 720 data-points. In contrast to the parent experiment, each participant conducted only one experiment.

Outcome Variables

For creating the 45 x 45 correlation matrix, five behavioral psychological and five physical variables were generated.

The five behavioral psychological variables were generated by the behavior of the participants and defined as follows:

T1: Number of left key presses
T2: Number of right key presses
T3: Number of double key presses
DR: Mean time between key presses, i.e., speed
DV: Mean variance between key presses, i.e., constancy of behavior

The five physical variables were associated with the behavior of the random event generator (REG) and derived from the following values:

TR: Number of times the output of the Markov-chain parsing of the REG yielded "1" during one run, i.e., the physical behavior of the REG filtered by the Markov chai
DT: The number of steps the fractal display deviated from the experimental instruction in either direction or from the central position, i.e., this is the summarized number of steps the fractal deviated from the goal
KR: deviation of the actual physical output of the Markov chain from an ideal Markov chain, measured as the deviation of the theoretical autocorrelation function from the experimental autocorrelation function of the sub-run calculated over 10 steps
ZT: mean voltage output of the REG at channel 4 out of eight; this channel was defined a priori as the one where the voltage would be recorded, because it was the middle channel and hence least likely to be affected by currency changes due to physical switching processes; the other channels were measured but the data not checked and analyzed
ZV: the standard deviation of this voltage output at channel 4, i.e., of the variable ZT.

There was continuous voltage applied to the Zener diode which triggered a current. This randomly changing current was converted by analogue–digital converter. Each time a key press was enacted the converter was sampled. If the number of bits was smaller than the previous one the outcome was 0, if it was larger the outcome was 1, and if it was equal a new sampling was initiated.

The 5 variables were calculated for each run per participant. As each participant had nine runs, this yielded 45 behavioral-psychological and 45 physical variables. These variables were correlated across all participants, which together produced the 45*45 matrix with 2025 cells. These cells were filled by the Spearman rank correlations coefficients between the respective variables across participants. We counted the number of correlations significant at the predetermined level of $p < .1$ (one-sided, or .05 two-sided). This is arbitrary and followed previous practice and our protocol. We also report sensitivity analyses for correlations significant at a lower p-value than that. The idea behind the testing procedure is as follows: In each correlation matrix there is a number of correlations significant at a certain level by chance. For instance, in a matrix of 100 cells there would be 5 correlations expected to be significant at the level $p = .05$ or 10 at the level $p = .1$. Similarly, in a matrix of 2025 cells we would expect 202 to 203 correlations to be significant at a level $p = .1$. Therefore, we counted the number of correlations significant at the level $p \le .1$ and tested (see below), whether this number of significant correlations found empirically was significantly different from a chance finding, using a randomization test, or different from the number of correlations found in a control experiment.

Control Experiments

After each participant had finished his or her experiment, the experimenter started a control experiment and then left the room. The control experiment consisted of the physical equipment running empty. This resulted in the generation and recording of the physical variables (TR, DT, KR, ZT, ZV as described above) without interference or interaction from a participant, sampling as many data points as during a real experiment. The generated array of physical variables was automatically written into a database, and the psychological variables of the previous experiment copied into the control database as corresponding psychological variables. Thus, each real experiment was matched by a control experiment with the same set of psychological variables, whereby all potential causal and non-causal effects were transferred into the control database and correlated with a new set of independent physical variables.

Special Features of the Two Experiments

Experiment 1, conducted by KK in Dresden, had, in addition, the following features: Instead of performing just one control experiment at the end of each session, a second control experiment was carried out at the beginning. In this way, the scope of the comparison was expanded. Furthermore, the questionnaire used in previous matrix experiments was modified by adding state variables identified as particularly psi-promoting (see Braud, 2002) and discussed by the matrix experiment consortium against the background of the Organizational Closure. These variables served as foundation of data to exchange the psychological variables (keystrokes) with the questionnaire data, using the “Phenomenology of Consciousness Inventory” (Pekala, 1995) in an additional analysis and to perform explorative analyses to examine which variables, determined as favoring psi effects, influence the number of significant correlations. The questionnaire was implemented in an online format on the computer where the experiment took place. In addition, this questionnaire was continued after the experiment had been performed in order to allow a comparison of the participants’ states before and after the experiment. The experimenter was blinded to the responses during the experiment; only after the experiment was completed were the answers inspected.

Experiment 2, conducted by HV in Witten used two additional features: There was a switch implemented that allowed the system to choose between two types of REGs. One was the custom-made REG that was also used by KK, identical to the ones from the first experiment. The second was an off-the-shelf REG called TrueRNG which can be easily purchased and implemented via a USB-stick. The idea was to see whether our elaborate sampling process would really be better or whether we might be able to offer a simpler system for wider usage. A coin toss at the beginning of the experiment decided which REG would operate the experiment. The second feature referred to the implementation of an assessment of absorption (Glicksohn, 2001; Glicksohn et al., 1992; Watt & Tierney, 2013). It consists of measuring objective time the experiment takes and then asking participants to estimate the time they took to conduct the experiment. The difference can serve as a measure of absorption, as more deeply absorbed participants tend to underestimate the time (Sedlmeier et al., 2020).

Ethical clearances were given by the respective ethical boards.

Statistical Analysis and Data Preparation

Data analysis followed the predefined protocol and consisted in a randomization test as specified. Briefly, an analysis script was written in Matlab to reshuffle the data 10,000 times and to recalculate the correlation coefficients each time (see Appendix). For every permutation step, the number of significant correlations in the matrix was counted. The number of times out of those 10,000 permutations where an equal or larger number of correlations was found than in the empirical matrix, divided by 10,000, yields an estimate of the true probability that the empirical result or a more extreme one could have been found by chance.

The experiment might be challenged to be open to systematic causal coding, for instance if someone used a certain strategy such as hammering on the keyboard, or always alternating shift keys, there might be causal correlations between physical and psychological variables. Therefore, we defined a sensitivity analysis: We analyzed only those correlations that are found in the time-forward or upper part of the matrix. As the matrix unfolds 9 * 5 psychological variables in rows and 9 * 5 variables in columns the correlation of the first set of physical variables with the second set of psychological variables is a time-forward correlation of physical variables in the first run with psychological variables in the second run, which should preclude all causality, as causality normally does not run backwards in time.

In the second experiment, the data for the TrueRNG in Experiment 2 were found to not conform to expected behavior (Appendix Figure 1 and Appendix Figure 2). Closer inspection revealed that this was due to a newly acquired programming glitch when programming the switch between the REGs that led to a buffer overflow for the data coming from the TrueRNG. We normalized the data and after normalization they conformed well to chance expectation (Appendix Figure 3 and Appendix Figure 4). The programming mistake was corrected for subsequent usage.

RESULTS

Experiment 1—Dresden Experiment by KK

Sixty-four participants were recruited, 43 females (67%) and 21 males (33%). Due to the layout of the questionnaire, which used the original one by Walter von Lucadou, age was only available in categories. The category of 41 to 50 years was the modal one with 17 participants. Two participants were below 20 years, six were below 30, and 14 were between 31 and 40. Sixteen participants were between 51 and 60, seven were between 61 and 70 and one person was older between 71 and 80.

The results of the statistical analysis of Experiment 1 can be seen in Table 1. (Appendix Table 1 presents the data together with the results of the control matrices.)

TABLE 1
Result of Statistical Analysis (Permutation Test with 10,000 Iterations) of Experimental Matrix, Full 45*45 Matrix, Experiment 1.
Yellow: Significant Results; Red: Missing Significance

45x45	**sig_th**	**0,1**	**0,05**	**0,02**	**0,01**	**0,005**	**0,002**	**0,001**	**0,0005**	**0,0002**	**0,0001**
full	**zo**	245,00	141,00	69,00	37,00	13,00	10,00	6,00	4,00	1,00	1,00
full	**n_sim**	1069	594	401	551	2588	611	559	390	963	417
full	**p_sim**	0,107	0,059	0,040	0,055	0,259	0,061	0,056	0,039	0,096	0,042
part	**zo_part**	105,00	44,00	17,00	9,00	1,00	1,00	0,00	0,00	0,00	0,00
part	**n_part_sim**	2023	4854	4793	4057	8230	4612	4874	3042	1447	809
part	**p_part_sim**	0,202	0,485	0,479	0,406	0,823	0,461	0,487	0,304	0,145	0,081

sig_th: theoretical significance level at which the number of significant correlations is counted
zo: number of significant correlations empirically found at respective level
n_sim: number of simulated matrices out of 10,000 with significant correlations at or above the number found empirically
p_sim: actual significance level of observed number of correlations (n_sim/10,000)
zo_part: number of correlations in time-forward (upper) part of the matrix
n_part_sim: number of significant correlations found in 10,000 simulations at respective level in upper part of the matrix
p_part_sim: actual significance level of observed number of correlations (n_part_sim/10,000) in upper part of the matrix

The first line of Table 1 presents the significance level at which the numbers of significant correlations are counted, the second line gives the empirically found number of significant correlations at that level. The number of simulated matrices with significant correlations at or above the number found empirically out of 10,000 simulations follows in the next line, and the p-level is given by this number divided by 10,000. The red color indicates which one of those statistical tests did not reach formal significance, while the yellow color indicates significance. The lower part of Table 1 reports the same for the upper diagonal of the correlation matrix, which contains only time-forward correlations, i.e., the correlation of the physical variables in the first run with the psychological variables of the second run (abbreviated as "part" in Table 1). This contains the causally independent parts of the correlation matrix because they are time-forward.

Only for some of the levels of significance were there more significant correlations than expected by chance (remember that $p \leq .1$ was the predefined level), namely for correlations at the level of $p \leq .02$, $p \leq .0005$, and $p \leq .0001$. The number of significant correlations at $p \leq .05$ and $p \leq .01$ miss formal significance by a small margin. The number of significant correlations at the predefined level of $p \leq .1$ is not significant.

While in the original experiment (Walach et al., 2020) we found significant correlations beyond chance even in the time-forward upper part of the matrix, overall none could be determined in this case.

We also analyzed smaller matrices (27*45, 18*27) which correspond to the setup of previous experiments by Walter von Lucadou and can be considered as replications of the earlier experiments. None of them showed any consistent and clear-cut results (Appendix Table 2 and Appendix Table 3).

Experiment 2—Witten Experiment by HV

The experiment conducted by HV in Witten recruited 40 participants, all of them students at the university and most of them psychology students who received course credits. Thus, all of them were between 18 and 30 years old. In that experiment we also measured time and had participants estimate the time of the experiment. On average, participants estimated the experiment as 0.4 minutes shorter than it actually was, which is a sign of modest absorption or closure. Twelve participants reckoned that the experiment took longer than it actually took, and thus were likely not very involved. The result of the statistical analysis is given in Table 2. Graphical representations of the experimental matrices of Experiments 1 (KK) and 2 (HV), as well as one of the control experiments (by KK) are presented in Figures 2, 3, and 4.

As can be seen, at none of the evaluated levels of significance do we find more significant correlations than expected by chance, neither in the full matrix, nor in the partial one. The same is true for the smaller matrices (27*45, 18*27 matrix; data not shown). Because there was no systematic effect in the first place, further analyses as to the efficacy of the two different REGs or the importance of organizational closure, measured as absorption were no longer useful. The control matrices did not show a significant effect either.

Taken together, none of the two experiments corroborates our original findings and the replication must be considered failed.

TABLE 2

Result of Statistical Analysis (Permutation Test with 10,000 Iterations) of Experimental Matrix, Full 45*45 Matrix, Experiment 2.

Yellow: Significant Results; Red: Missing Significance

45x45	**sig_th**	**0,1**	**0,05**	**0,02**	**0,01**	**0,005**	**0,002**	**0,001**	**0,0005**	**0,0002**	**0,0001**
full	**zo**	160,00	89,00	35,00	18,00	7,00	2,00	0,00	0,00	0,00	0,00
full	**n_sim**	9437	7347	6865	5981	7219	7206	8287	6162	3395	1966
full	**p_sim**	0,945	0,735	0,687	0,598	0,722	0,721	0,829	0,616	0,340	0,197
part	**zo_part**	79,00	46,00	19,00	12,00	4,00	0,00	0,00	0,00	0,00	0,00
part	**n_part_sim**	7329	4308	3963	2234	4702	7848	5587	3556	1700	927
part	**p_part_sim**	0,733	0,431	0,396	0,223	0,470	0,785	0,559	0,356	0,170	0,093

sig_th: theoretical significance level at which the number of significant correlations is counted
zo: number of significant correlations empirically found at respective level
n_sim: number of simulated matrices out of 10,000 with significant correlations at or above the number found empirically
p_sim: actual significance level of observed number of correlations (n_sim/10,000)
zo_part: number of correlations in time-forward (upper) part of the matrix
n_part_sim: number of significant correlations found in 10,000 simulations at respective level

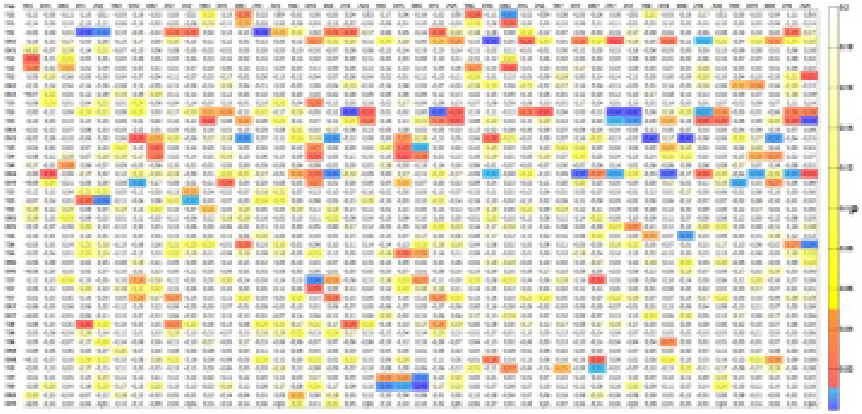


Figure 2. Experimental Matrix of Experiment 1 (KK).

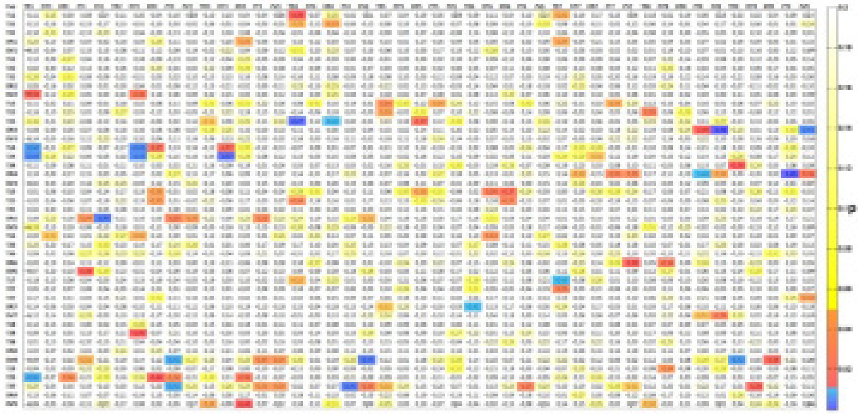


Figure 3. Experimental Matrix of Experiment 2 (HV).

Figure 4. Control Matrix of Experiment 1 (KK).

DISCUSSION

Our hope that we might be able to replicate our earlier positive finding and those reported by Walter von Lucadou (Lucadou, 1986, 1987b, 1991, 2000; Lucadou et al., 1987; Walach et al., 2020) did not bear out. The two experiments, reported here, were part of a concerted effort to find a replicable experimental model that would circumvent the NT axiom. This prohibits signal coding for anomalous experiments which are supposed to operate on the basis of generalized entanglement correlations. Circumventing the NT axiom was not possible. By the same token, our negative results also preclude an anomalous signal. For had such an anomalous signal been there, we would have been able to see it, as it would have driven one of our variables (TR) that measures the deviation of the REG from randomness, and thus produced a series of significant correlations. This result has to be seen against potential weaknesses and against other results, partially positive and partially negative.

A major weakness of our experiments is that they are comparatively small. So, one could argue that they did not have the necessary power. While our predecessor experiment had 503 participants, these new experiments only had 104 participants together. Using the effect size of our predecessor experiment, approximately r = .38, our smaller experiment had a power of 69% and the larger one a power of 88% to detect the effect. Precisely because we assume that the effects are of a non-classical, non-signal–like nature a classical power discussion is beside the point, we contend. A classical power analysis assumes that there is a stable effect that can be detected, given enough resources. We do not think that this is the case. This is the reason why in our consortium protocol power analysis is not part of the protocol, but only definition of recruitment procedures and a preclusion of optional stopping. As we argue below, power is not the decisive issue, as there are various instances of strongly powered and well-prepared replications that were unsuccessful.

This lack of success in replication is not a problem of personal factors, as these experiments were conducted by two independent groups following the same protocol and using the same equipment. Rather, it feeds into a stream of similar results: Walleczek and von Stillfried (2019) were unable to replicate the Radin double-slit experiment, a careful replication in which Radin himself was involved, conducted the experiment according to a predefined protocol, and analyzed the data according to previous standards. Rabeyron was unable to replicate Bem's retro-priming results (Rabeyron, 2020). Maier and colleagues were unable to replicate earlier results and found higher-level regularities, i.e., an effect that moves in a kind of sinusoidal wave from positivity to negativity and potentially back again (Dechamps & Maier, 2020; Maier & Dechamps, 2018; Maier et al., 2018). The matrix experiment was repeated in a different form by Grote, who could not find clear-cut effects either (Grote, 2015, 2017). A newly designed experiment by Grote which replicated the general setup of the matrix experiment with new equipment and 200 participants was unsuccessful. This demonstrates that power does not seem to be the issue. However, an analysis of correlations of the same physical data of this experiment with different psychological variables, in that case questionnaire data obtained from each participant before the experiment, was marginally significant ($p = 0.064$) (Grote, 2021). Jolij and Bierman conducted two replications of Bem's retropriming paradigm, but found no effect (Jolij & Bierman, 2019). However, when they analyzed the questionnaire data that were also taken together with the psi data in a matrix analytical approach, they found a significant result ($p < 0.03$) in one experiment, the smaller one with 61 participants, and a borderline significant effect in the second study ($p = 0.06$) with 222 participants. This is again a clear hint that the decisive question is not about power.

These results have to be seen together with experiments by Ana Borges in Edinburgh who has conducted three experiments herself with clearly positive results and one commissioned by another experimenter with negative results (Ana Borges, personal communication and unpublished Ph.D. thesis, The University of Edinburgh, Department of Psychology).

The results of the experiments of Ana Borges can only be really discussed once they are fully published. Meanwhile one might suppose that in those experiments we are dealing with an experimenter effect, as the study conducted by a second experimenter who was indifferent to the results was clearly negative, while the studies conducted by Borges herself, who is enthusiastic about this work, were positive. We had such a setup implicitly in our experiments: KK tended toward hoping to find positive results, while HV was pretty indifferent toward the results of the experiment. Was the negative effect of HVs experiment a negative experimenter effect? As the CIRTS theory would suggest, all experiments might be in principle tests of experimenter PK (or the lack of it) (Bierman, 2008).

A strength and weakness of our experiment at the same time was the statistical analysis. The Monte-Carlo simulation of potential different matrices produces an empirical distribution against which statistical inferences can be made without any parametric assumptions and is thus a straightforward, non-parametric analysis. It is comparatively stable: The p-values change maximally by 10^{-3} and the values of significant simulated matrices by around 30, if the 10,000 iterations are repeated 30 times, i.e., instead of 1,069 significant matrices which translates into $p = 0.106$ we would have 1,099 significant matrices which translates into $p = 0.109$. It also corrects for potential causal biases, as these are destroyed in the permutations.

But such an analysis also destroys the intricate network between potential causal and non-local correlations, making the analysis conservative. The type of analysis chosen and defined in the protocol actually uses only the experimental matrix. One could also use difference scores between the experimental and the control matrix and other metrics for the statistical analysis. We have done that for exploratory purposes. But this does not change the result.

An optimal analysis might be able to use some difference metric between the control and the experimental matrix. One might argue that the effect is embedded within the whole experiment and not only within the experimental matrix. Thus, some difference measure between the two matrices might be better able to capture the effect. This is for a subsequent analysis of the data to decide.

In our view, the results seem to suggest a decline effect as observed by Maier and colleagues: Our own first experimental results were the stimulus for further work. They were very positive. The experiments of Ana Borges were immediate successor experiments timewise and were also positive. KK's experiment was next and had a small, nearly significant effect. HVs experiment was the last in this series and had a zero effect. This supports a decline effect and contradicts our expectation that the matrix method might help to mitigate such a decline. A decline effect is a prediction of our model (Lucadou, 2015b; Lucadou et al., 2007; Walach et al., 2014): The NT axiom states that whenever effects due to generalized entanglement correlations are mistaken as causal effects and could be used for signal transmission. the effects go away (decline), or change channel, i.e., become visible in another parameter not tested, or change sign, i.e., become obvious in the control group.

Obviously, the NT axiom (Lucadou et al., 2007) cannot be circumvented as we had hoped. It may take longer before a decline comes into effect. But eventually there is no experimental system that generates its own comparison standard through a control group that can elude it. For no matter how complex the system or how many degrees of freedom, eventually there will always be an option to code a signal. In our case it would have been the number of significant correlations.

We had similar experiences with other experimental models. A careful pilot study of a DMILS replication, in which we tried to replicate the originally successful DMILS studies of Schlitz and Braud (Braud & Schlitz, 1983; Braud & Schlitz, 1991; Schlitz & Braud, 1997), yielded a strong positive effect of r = .35, which was, however, not tested statistically as per protocol (Schmidt et al., 2001). A large replication with sufficient participants for detecting a much smaller effect failed utterly (Schmidt, 2002; Schmidt et al., 2002). We replicated the Grinberg-Zylberbaum study in which he had claimed that a visual stimulation of one subject had introduced transferred evoked potentials in the EEG of a spatially distant, but connected participant (Grinberg-Zylberbaum et al., 1994). In our study we could not find transferred potentials as such, but significant deviations from chance expectations (Wackermann et al., 2003). Harald Walach commissioned two large-scale replications in the same lab, which were clearly positive, but never published (Claudio Naranjo, personal communication; he had conducted the studies but was prohibited from publishing the data by Wackermann, the former head of the lab). We thought we had a replicable, if complicated paradigm and conducted another replication which was meant to be completely foolproof against fraud and artifacts, as it was between subjects separated by about 800 kms. But we could not find the effect in its original signature. We found an effect in the alpha frequency band which was significant in three studies. However, the relevance of this effect remains unclear as it only showed up after averaging thousands of trials. Instead, we saw an unexpected anticipatory or precognition effect (Hinterberger et al., 2008, 2007). The reverse priming study by Daryl Bem (Bem, 2011) did not prove to be as replicable as hoped either (Jolij & Bierman, 2019; Rabeyron, 2014; Ritchie et al., 2012).

It seems we have enough controversial data and failed replications. It is important to note at this point: Failed replications and positive results in meta-analyses do not contradict each other. It is possible that in a long series of experiments some very careful negative replications, although they might be important, either do not (Schmidt et al., 2004b), or only partially (Bösch et al., 2006) influence the summary result of the meta-analysis, because many other positive results are published or because effects that have been negative in the hands of one research group recover in other labs (Bierman, 2001). This is to be expected under the NT axiom, since it only applies to strict replications. As soon as parameters are changed, and they usually are when other groups replicate an experiment, it is, technically speaking, a new experiment, even though it might use the same experimental model and will be analyzed under the same umbrella by meta-analysts. Thus, one way out of the conundrum would be to conduct replications as conceptual replications, changing important elements in an experimental paradigm so as to prevent it from being a direct replication which could be used for signal coding.

Another thought might be worth considering: If our hypothesis is correct and generalized entanglement correlations exist and are the basis for most, if not all PSI phenomena, then we need to consider the fact that in real life they are normally always embedded in a series of local-causal correlations which also support and frame them, like water is supported in a sponge (Lucadou, 2019). In the experimental situation we are trying to separate the two out, squeezing the sponge, as it were, and then are surprised to find the structure and the water gone.

Thus, the current situation is an impasse: The directly replicable paradigm that critics demand seems to be impossible. The fact that so many studies have been conducted by different groups and in slightly varying designs allows meta-analysts to draw positive conclusions. Hence, both skeptics and proponents of PSI are right and wrong at the same time. The "Dodo bird verdict" which has beset psychotherapy research is valid here as well: All have won and all must have prizes (Luborsky et al., 2002; Rosenzweig, 1936). It has been pointed out that this constitutes a paradox: If PSI is real, as a lot of the data suggest, then by the same token it cannot be proven experimentally, because the experimental paradigm presupposes the possibility of partitioning reality into independent segments, which is exactly what PSI negates (Rabeyron, 2020).

What this series of replications together with other evidence shows, is in our view that a causal, signal-theoretical interpretation of PSI is unlikely. It rather strengthens, even though indirectly, an analysis and theoretical model that assumes these effects to be instances of generalized entanglement correlations, or similar processes. If so, critics will remark: Why is it that entanglement correlations could be empirically proven in the physical case, but not in such a generalized case as in parapsychology? The answer to this question is straightforward: In the physical case we have a very strong formalism that allows the derivation of expectation values or empirical bounds that are theoretically defined, such as Bell's inequalities. This defined frame is not given in the generalized case because the model is not strong enough and does not contain enough quantitative terms that would allow such a derivation. In the physical case, only combinations of for example polarization angles are measured, and whether they are correlated or not is *not* determined by an experimental control group of different or incompatible angles, but by the violation of Bell's inequalities, i.e., by the theoretical distribution of two joint probabilities. This is structurally completely different from determining the control standard by a control experiment. As long as we do not have an equally strong theoretical framework, we will not be able to provide a straightforward proof of the facticity of generalized entanglement correlations.

Proponents of remote viewing experiments often lament about the inadequacy of experimenting with people who have no special gift for PSI, as is the rule in experiments like ours (May et al., 2018). They liken it to trying to judge musical prowess in an average group of people, some of whom might be musically gifted while the majority won't be, diluting the end result. Experimenting with gifted people might help avoid this pitfall. However, it was estimated that this will be maximally one or two in one hundred (May et al., 2018). While this argument is certainly convincing in part, it conflates two distinct points: Working with gifted people is certainly a good idea. But this does not preclude failures, as the failed replication by Walleczek and von Stillfried (2019) showed. The remote viewing experiment is not an experiment in the sense the term is used here. And this might be the reason why remote viewing experiments cannot violate the NT axiom and hence can produce quite stable results (Targ, 2019; Targ & Katra, 2000).

In remote viewing there is no control standard that is *produced by the experiment*. The control is the expectation of no special information transferred, which is a generic null-expectation. Therefore, it can be replicated at will. The NT axiom would only come into effect in the counterfactual situation, which by definition never exists, if the same person were to target the very same target twice. But the same remote viewer will not normally do this, and once a target is described there is no point in having this repeated. Also, in experimental setups that are similar, targets and participants are normally changed, thus implicitly avoiding the NT axiom. Therefore, some free-response remote viewing or telepathy studies might be able to eschew the NT axiom, but all studies that produce their own control standard in a control group and are replicated as an exact replication will have the same problems as we experienced. Unfortunately, remote viewing and Ganzfeld telepathy studies belong to a category where a lot of expert knowledge, material, and facilities are necessary and hence do not lend themselves to the type of classroom experiment that is set up quickly and easily to demonstrate telepathy.

Thus, we might have to live with the fact that a definitive experimental paradigm is very difficult, if not impossible, to have. As long as a paradigm incorporates enough changes, for instance by way of conceptual replication, or changing variables, or outcome measures each time it is conducted, it may eschew the NT axiom. But by the same token it will also be less convincing to skeptics, who will keep demanding a strict replication. Thus, skeptics will likely have an easy life: They won't be bullied into acceptance by a foolproof experimental paradigm of PSI, because it simply may not exist. So, is experimenting, then, unnecessary and a waste of time and resources? Probably not, because it might teach us about higher order parameters, such as the recovery time it takes until an effect bounces back, or about the amount of change necessary to make an experiment conceptually a new one (Dechamps & Maier, 2020; Maier et al., 2018). Or it might help decide between theoretical options (Bierman, 2010). Or it might yield a higher class of models that not only predict when an effect might appear, but also when it will go away. Experimenting might thus also produce the parameters necessary to build a fuller model that contains enough richness to derive a formally more stringent theory.

But we should probably give up the hope that the intellectual fight about whether anomalous cognition effects or PSI is real, can be won with the brute force of rational argument and experimental evidence alone. This is very rarely the final arbiter anyway, even for very mundane questions, where social movements, intellectual fashions, generic worldviews, political considerations are often much more important (Latour, 1999). Perhaps a mixed approach will be best: devising clever experiments, avoiding the pitfalls of the NT axiom by changing procedures in replications, not forgetting qualitative real-world studies, observations of natural occurrence of PSI and analytical arguments combating the prevailing naturalistic stance that is more of a dogma than an intellectual necessity (van Fraassen, 2016; Williams & Robinson, 2016). All this together might help opening up the community for the possibility of PSI. Producing a final proof is likely a vain expectation, as our results show.

Our conclusion is: The matrix experiment is likely not a replicable experiment. The NT axiom that prohibits signal transfer in systems that are built on correlations might be operative even in this sophisticated experimental design. This makes likely that such effects are not of a local-causal nature. In addition, artefacts might be operative in this highly complex study. There might be other regularities involved which we do not understand as yet, but we can preclude signals with a high likelihood, else we would have seen their effect. Future studies should determine if conceptual replications of the matrix experiment changing important elements and parameters can avoid the NT axiom. In addition, further research efforts could advance the experimental setup of the matrix experiment/CMM, transferring it to other psi areas.

ACKNOWLEDGMENTS

This work was funded by Bial Grant 400/14. The study by KK was a master's thesis at the Technical University Chemnitz. We thank Nikolaus von Stillfried for important conceptual and logistic help, especially with the expert symposium that allowed the formulation of the consensus protocol. We thank Torkel Falkenberg who helped secure co-funding for the expert-symposium and the meeting venue. We are grateful to all the experts who participated and who helped formulate this protocol in various Delphi rounds.

Author Statements: HW developed the original protocol, secured funding, organized Experiment 2, and wrote the first draft of the manuscript. He participated in discussing and finalizing the manuscript. KK and HV conducted the experiments and collected data. PS supervised Experiment 1. WvL was the senior advisor in this project, provided his material, and helped with the understanding of the experimental procedures and discussion of results. TH conducted the statistical analyses. All authors participated in writing, reviewing, and approving the final manuscript.

Conflict of Interests: None of the authors has a conflict of interest.

REFERENCES

Abbot, B. P., & LIGO Scientific Collaboration and Virgo Collaboration. (2016). Observation of gravitational waves from a binary black hole merger. *Physical Review Letter, 116*(061102). doi:10.1103/PhysRevLett.116.061102

Alcock, J. E. (2003). Give the null hypothesis a chance: Reasons to remain doubtful about the existence of PSI. *Journal of Consciousness Studies, 10*(6–7), 29–50.

Aspect, A., Dalibard, J., & Roger, G. (1982a). Experimental test of Bell's inequalities using time varying analyzers. *Physics Review Letter, 49*, 1804–1807.

Aspect, A., Grangier, P., & Roger, G. (1982b). Experimental realization of Einstein-Podolsky–Rosen–Bohm–Gedankenexperiment: A new violation of Bell's inequalities. *Physics Review Letter, 49*, 91–94.

Atmanspacher, H., & Primas, H. (2006). Pauli's ideas on mind and matter in the context of contemporary science. *Journal of Consciousness Studies, 13*(3), 5–50.

Atmanspacher, H., Primas, H., & Wertenschlag-Birkhäuser, E. (Eds.). (1995a). *Der Pauli-Jung-Dialog und seine Bedeutung für die moderne Wissenschaft*. Springer.

Atmanspacher, H., Römer, H., & Walach, H. (2002). Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of Physics, 32*, 379–406.

Atmanspacher, H., Wiedenmann, G., & Amann, A. (1995b, January/February). Descartes revisited: The endo-exo distinction and its relevance for the study of complex systems. *Complexity*. https://onlinelibrary.wiley.com/doi/10.1002/cplx.6130010306

Bell, J. S. (1987). *Speakable and unspeakable in quantum mechanics*. Cambridge University Press.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100*, 407–425.

Bierman, D. J. (2000). On the nature of anomalous phenomena: Another reality between the world of subjective consciousness and the objective world of physics? In P. van Loocke (Ed.), *The physical nature of consciousness* (pp. 269–292). John Benjamins.

Bierman, D. J. (2001). On the nature of anomalous phenomena: Another reality between the world of subjective consciousness and the objective world of physics? In P. Van Loocke (Ed.), *The physical nature of consciousness* (pp. 269–292). Benjamins.

Bierman, D. J. (2008). *Consciousness induced restoration of time symmetry (CIRTS), a psychophysical theoretical perspective*. Paper presented at the Parapsychological Association 51st Annual Convention, Winchester, UK.

Bierman, D. J. (2010). Consciousness induced restoration of time symmetry (CIRTS): A psychophysical theoretical perspective. *Journal of Parapsychology, 74*(2), 273–299.

Bösch, H., Steinkamp, F., & Boller, E. (2006). Examining psychokinesis: The interaction of human intention with random number generators—A meta-analysis. *Psychological Bulletin, 132*, 497–523.

Braud, W. (2002). Psi-favorable conditions. In V. G. Rammohan (Ed.), *New frontiers of human science: A festschrift for K. Ramakrishna Rao* (pp. 95–118). McFarland.

Braud, W., & Schlitz, M. (1983). Psychokinetic influence on electrodermal activity. *Journal of Parapsycholoty, 47*, 95–119.

Braud, W., & Schlitz, M. J. (1991). Conscious interactions with remote biological systems: Anomalous intentionality effects. *Subtle Energies, 2*, 1–46.

Braude, S. E. (1986). *The limits of influence. Psychokinesis and the philosophy of science*. Routledge and Kegan Paul.

Braude, S. E. (2017). The mediumship of Carlos Mirabelli (1889–1951). *Journal of Scientific Exploration, 31*, 435–456.

Cardeña, E. (2018). The experimental evidence for parapsychological phenomena:

A review. *American Psychologist, 73*(5), 663–677. doi:10.1037/amp0000236

Carr, B. J. (2015a). Higher dimensions of space and time and their implications for psi. In E. May & S. Marwaha (Eds.), *Extrasensory perception: Support, skepticim and science, Vol. 2* (pp. 21–61). Greenwood.

Carr, B. J. (2015b). Hyperspatial models of matter and mind. In E. F. Kelly, A. Crabtree, & P. Marshall (Eds.), *Beyond physicalism: Toward reconciliation of science and spirituality* (pp. 227–273). Rowman & Littlefield.

d'Espagnat, B. (1997). Aiming at describing empirical reality. In R. S. Cohen, M. Horne, & J. Stachel (Eds.), *Potentiality, entanglement, and passion-at-a-distance* (pp. 71–87). Kluwer.

Dechamps, M. C., & Maier, M. A. (2020). How smokers change their world and how the world responds: Testing the oscillatory nature of micro-psychokinetic observer effects on addiction-related stimuli. *Journal of Scientific Exploration, 33*, 406–434. doi:10.31275/2019/1513

Dobyns, Y. H., & Nelson, R. D. (1998). Empirical evidence against decision augmentation theory. *Journal of Scientific Exploration, 12*, 231–257.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of reality be considered complete? *Physical Review, 47*, 777–780.

Filk, T., & Römer, H. (2011). Generalized quantum theory: Overview and latest developments. *Axiomathes, 21*, 211–220. doi: 10.1007/s10516-010-9136-6

Glicksohn, J. (2001). Temporal cognition and the phenomenology of time: A multiplicative function for apparent duration. *Consciousness and Cognition, 10*, 1–25.

Glicksohn, J., Mourad, B., & Pavell, E. (1992). Imagination, absorption and subjective time estimation. *Imagination, Cognition and Personality, 11*(2), 167–176.

Grinberg-Zylberbaum, J., Delaflor, M., Attie, L., & Goswami, A. (1994). The Einstein-Podolsky-Rosen paradox in the brain: The transferred potential. *Physics Essays, 7*, 422–427.

Grosso, M. (2016). *The man who could fly: St.Joseph of Copertino and the mystery of levitation*. Rowman & Littlefield.

Grote, H. (2015). A correlation study between human intention and the output of a binary random event generator. *Journal of Scientific Exploration, 29*, 265–290.

Grote, H. (2017). Multiple-analysis correlation study between human psychological variables and binary random events. *Journal of Scientific Exploration, 31*, 231–254.

Grote, H. (2018). *Gravitationswellen. Geschichte einer Jahrhundertentdeckung [Gravitation waves. History of a once-in-a-century discovery]*. Beck.

Grote, H. (2021). Mind-matter entanglement correlations: Blind analysis of a new correlation matrix experiment. *Journal of Scientific Exploration*, 35(2), 287–310.

Hameroff, S., & Penrose, R. (2014). Consciousness in the universe: A review of the 'Orch OR' theory. *Physics in Life Review, 11*, 39–78.

Handsteiner, J., Friedman, A. S., Rauch, D., Gallicchio, J., Liu, B., Hosp, H., . . . Zeilinger, A. (2017). Cosmic Bell test: Measurement settings from Milky Way stars. *Physical Review Letters, 118*(6), 060401-060401-060408. doi:10.1103/PhysRevLett.118.060401

Heim, B. (1989). *Elementarstrukturen der Materie. Einheitliche strukturelle Quantenfeldtheorie der Materie und Gravitation Bd. 1*. Resch-Verlag.

Hinterberger, T., Mochty, U., Schmidt, S., Erat, L.-M., & Walach, H. (2008). EEG Korrelationen zwischen räumlich weit entfernten Paaren [EEG correlations between spatially distant pairs]. *Zeitschrift für Anomalistik, 8*, 55–75.

Hinterberger, T., Studer, P., Jäger, M., Haverty-Stacke, C., & Walach, H. (2007). The slide-show presentiment effect discovered in brain electrical activity. *Journal of the Society of Psychical Research, 71*, 148–166.

Horton, R. (2015). Offline: What is medicine's 5 sigma? *Lancet, 385*, 1380. doi:10.1016/S0140-6736(15)60696-1

Houtkooper, J. M. (2002). Arguing for an observational theory of paranormal phenomena. *Journal of Scientific Exploration, 16*, 171–186.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124.

Jahn, R. G., & Dunne, B. J. (1987). *Margins of reality. The role of consciousness in the physical world*. Harcourt Brace Jovanovich.

Jahn, R. G., Dunne, B. J., Bradish, G. J., Dobyns, Y. H., Lettieri, A., Nelson, R. D., . . . Walter, B. (2000). Mind/machine interaction consortium: PortREG replication experiments. *Journal of Scientific Exploration, 14*, 499–555.

Jolij, J., & Bierman, D. J. (2019). Two attempted retro-priming replications show theory-relevant anomalous connectivity. *Journal of Scientific Exploration, 33*(1), 43–60.

Jung, C. G. (1952). Synchronizität als ein Prinzip akausaler Zusammenhänge. In C. G. Jung & W. Pauli (Eds.), *Naturerklärung und psyche* (pp. 1–107). Rascher.

Latour, B. (1999). *Pandora's hope: An essay on the reality of science studies*. Harvard University Press.

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., . . . Krause, E. D. (2002). The Dodo bird verdict is alive and well—Mostly. *Clinical Psychology: Science and Practice, 9*, 2–12.

Lucadou, W. v. (1974). Zum parapsychologischen Experiment—Eine methodologische Skizze. *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 16*, 57–62.

Lucadou, W. v. (1986). Keine Spur von Psi—Zusammenfassende Darstellung eines umfangreichen Psychokineseexperiments. *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 29*, 169–197.

Lucadou, W. v. (1987a). A multivariate PK experiment. Part I. An approach combining physical and psychological conditions of the PK process. *European Journal of Parapsychology, 6*(4), 305–345.

Lucadou, W. v. (1987b). A multivariate PK experiment. Part III. Is PK a real force? The results and their interpretation. *European Journal of Parapsychology, 6*(4), 369-428.

Lucadou, W. v. (1991). *Locating Psi-bursts—Correlations between psychological characteristics of observers and observed quantum physical fluctuations.* Paper presented at the The Parapsychological Association 34th Annual Convention, Proceedings of Presented Papers. Heidelberg.

Lucadou, W. v. (1995). The model of pragmatic information (MPI). *European Journal of Parapsychology, 11*, 58–75.

Lucadou, W. v. (2000). *Backward causation and the Hausdorff-Dimension of singular events.* Paper presented at the Proceedings of Presented Papers, The Parapsychological Association 43rd Annual Convention August 17–20, Freiburg.

Lucadou, W. v. (2006). Self-organization of temporal structures—A possible solution for the intervention problem. In D. P. Sheehan (Ed.), *Frontiers of time. Retrocausation—Experiment and theory* (pp. 293-315). American Institute of Parapsychology.

Lucadou, W. v. (2015a, July 16-19). *The correlation-matrix method (CMM)—A new light upon the repeatability problem of parapsychology.* Paper presented at the 58th Annual Convention of the Parapsychological Association, University of Greenwich.

Lucadou, W. v. (2015b). The Model of Pragmatic Information (MPI). In E. C. May & S. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science: Vol. 2: Theories and the future of the field* (pp. 221–242). Praeger.

Lucadou, W. v. (2019). Homeopathy and the action of meaning: A theoretical approach. *Journal of Scientific Exploration, 33*, 213–254. doi:10.31275/2019.1343

Lucadou, W. v., Lay, B., & Kunzmann, H. (1987). A multivariate PK experiment. Part II. Relationships between psychological variables. *European Journal of Parapsychology, 6*(4), 347–368.

Lucadou, W. v., Römer, H., & Walach, H. (2007). Synchronistic phenomena as entanglement correlations in generalized quantum theory. *Journal of Consciousness Studies, 14*(4), 50–74.

Ma, X.-S., Zotter, S., Kofler, J., Ursin, R., Jennewein, T., Brukner, C., & Zeilinger, A. (2012). Experimental delayed-choice entanglement swapping. *Nature Physics, online.* doi:10.1038/NPHYS2294

Maier, M. A., Büchner, V. L., Kuhbandner, C., Pflitsch, M., Fernández-Capo, M., & Gámiz-Sanfeliu, M. (2014). Feeling the future again: Retroactive avoidance of negative stimuli. *Journal of Consciousness Studies, 21*(9-10), 121–152.

Maier, M. A., & Dechamps, M. C. (2018). Observer effects on quantum randomness: Testing micro-psychokinetic effects of smokers on addiction-related stimuli. *Journal of Scientific Exploration, 32*, 265–297.

Maier, M. A., Dechamps, M. C., & Pflitsch, M. (2018). Intentional observer effects on quantum randomness: A Bayesian analysis reveals evidence against micro-psychokinesis. *Frontiers in Psychology, 9*(379). doi:10.3389/fpsyg.2018.00379

Mansfield, V. (1995). *Synchronicity, science, and soul-making. Understanding Jungian synchronicity through physics, Buddhism, and philosophy*. Open Court.

May, E. C., & Marwaha, S. B. (Eds.). (2018). *The Star Gate Archives. Reports of the United States Government sponsored psi program, 1972–1995. Vol 1: Remote viewing, 1972-1984*. Mc Farland.

May, E. C., & Marwaha, S. B. (Eds.). (2019a). *The Star Gate Archives. Reports of the United States Government-sponsored psi program 1972–1995. Vol. 3: Psychokinesis*. McFarland.

May, E. C., & Marwaha, S. B. (Eds.). (2019b). *The Star Gate Archives. Reports of the United States Government sponsored psi program, 1972–1995. Vol 2: Remote viewing: 1985–1995*. McFarland.

May, E. C., Spottiswoode, S. J., Utts, J. M., & James, C. L. (1995). Applications of Decision Augmentation Theory. *Journal of Parapsychology, 59*, 221–250.

May, E. C., Spottiswoode, S. J. P., & Faith, L. V. (2000). The correlation of the gradient of shannon entropy and anomalous cognition: Toward an AC sensory system. *Journal of Scientific Exploration, 14*, 53–72.

May, E. C., Utts, J. M., & Spottiswoode, S. J. P. (1996). Deicision augmentation theory: Applications to the random number generator. *Journal of Scientific Exploration, 9*, 453–488.

May, E. C., Utts, J. M., Trask, V. V., Luke, W. W., Frivold, T. J., & Humphrey, B. S. (2018). Review of the psychoenergetic research conducted at SRI International (1973–1988). In E. C. May & S. B. Marwala (Eds.), *The Star Gate Archives. Reports of the United States Government sponsored psi program, 1972-1995. Vol 1: Remote viewing, 1972–1984* (pp. 495-504). McFarland.

Meier, C. A. (Ed.) (1992). *Wolfgang Pauli und C. G. Jung. Ein Briefwechsel 1932–1958*. Springer.

Meier, C. A. (Ed.) (2001). *Atom and archetype: The Pauli/Jung letters 1932–1958*. Princeton University Press.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., . . . Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*, 0021. doi:10.1038/s41562-016-0021

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. doi:10.1126/science.aac4716

Pallikari, F. (2001). A study of the fractal character in electronic noise process. *Chaos, Solitons and Fractals, 12*, 1499–1507.

Peat, F. D. (1992). *Synchronizität: Die verborgene Ordnung [Synchronicity: The hidden order]*. Goldmann.

Pekala, R. J. (1995). A short, unobtrusive hypnotic-assessment procedure for assessing hypnotizability level: I. Development and research. *American Journal of Clinical Hypnosis, 37*, 271–283.

Penrose, R. (2004). *The Road to reality: A complete guide to the laws of the universe.* Jonathan Cape.

Primas, H. (1994a). Endo- and exo-theories of matter. In H. Atmanspacher & G. J. Dalenoort (Eds.), *Inside versus outside* (Vol. 63, pp. 163–193). Springer.

Primas, H. (1994b). Realism and quantum mechanics. In D. Prawitz, B. Skyrms, & D. Westerstahl (Eds.), *Logic, methodology and philosophy of science IX: Proceedings of the Ninth International Contress of Logic, Methodology and Philosophy of Science, Uppsala, Sweden, August 7-14, 1991* (pp. 609–631). Elsevier.

Rabeyron, T. (2014). Retro-priming, priming, and double testing: Psi and replication in a test–retest design. *Frontiers in Human Neuroscience, 8*(154). doi:10.3389/fnhum.2014.00154

Rabeyron, T. (2020). Why most research findings about psi are false: The replicability crisis, the psi paradox and the myth of Sisyphus. *Frontiers in Psychology, 11*(2468). doi:10.3389/fpsyg.2020.562992

Radin, D. (2018). *Real magic. Ancient wisdom, modern science, and a guide to the secret power of the universe*. Harmony.

Radin, D., Michel, L., Galdamez, K., Wendland, P., Rickenbach, R., & Delorme, A. (2012). Consciousness and the double-slit interference pattern: Six experiments. *Physics Essays, 25*, 157–171.

Reber, A. S., & Alcock, J. E. (2020). Searching for the impossible: Parapsychology's elusive quest. *American Psychologist, 75*, 391–399. doi:10.1037/amp0000486

Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's 'retroactive facilitation of recall' effect. *PLoS One, 7*(3), e33423. doi: 10/1371/journal.pone.0033423

Rosenzweig, S. (1936). Some implicit common factors in diverse methods in psychotherapy. *American Journal of Orthopsychiatry, 6*, 412–415.

Salart, D., Baas, A., Branciard, C., Gisin, N., & Zbinden, H. (2008). Testing spooky actions at a distance. *Nature, 454*, 861–864

Schlitz, M., & Braud, W. (1997). Distant intentionality and healing: Assessing the evidence. *Alternative Therapies in Health and Medicine, 3*, 38–53.

Schmidt, S. (2002). *Aussergewöhnliche Kommunikation? Eine kritische Evaluatgion* *des parapsychologischen Standardexperimentes zur direkten mentalen Interaktion*. BIS.

Schmidt, S. (2016). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 581–596). American Psychological Association.

Schmidt, S., Müller, S., & Walach, H. (2004a). *Do you know who is on the phone? Replication of an experiment on telephone telepathy*. Paper presented at the The Parapsychological Association 47th Annual Convention, Vienna.

Schmidt, S., Schneider, R., Binder, M., Bürkle, D., & Walach, H. (2001). Investigating methodological issues in EDA-DMILS: Results from a pilot study. *Journal of Parapsychology, 65*, 59–82.

Schmidt, S., Schneider, R., Utts, J., & Walach, H. (2002). Remote intention on electrodermal activity - Two meta-analyses. *Parapsychological Association Annual Convention*.

Schmidt, S., Schneider, R., Utts, J., & Walach, H. (2004b). Remote intention on electrodermal activity—Two meta-analyses. *British Journal of Psychology, 95*, 235–247.

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature, 470*, 437. doi:10.1038/470437a

Schrödinger, E. (1935). Discussion of probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society, 31*, 555–563.

Schwartz, G. E., Woollacott, M., Schwartz, S., Barušs, I., Beauregard, M., Dossey, L., . . . Tart, C. (2018). The Academy for the Advancement of Postmaterialist Sciences: Integrating Consciousness into Mainstream Science,. *Explore. The Journal of Science and Healing, 14*(2), 111–113.

Sedlmeier, P., Winkler, I., & Lukina, A. (2020). How long did the time spent in meditation feel? "Attention. Attention. Attention.". *Psychology of Consciousness: Theory, Research, and Practice*. https://doi.org/10.1037/cns0000254.

Sheldrake, R., & Smart, P. (2003). Videotaped experiments on telephone telepathy. *Journal of Parapsychology, 67*, 187–206.

Shimony, A. (1989). Search for a worldview which can accomodate our knowledge of microphysics. In J. T. Cushing & E. Mcmullin (Eds.), *Philosophical consequences of quantum theory: Reflections on Bell's Theorem* (pp. 25–37). University of Notre Dame Press.

Stefanov, A., Zbinden, H., Gisin, N., & Suarez, A. (2002). Quantum correlations with spacelike separated beam splitters in motion: Experimental test of multisimultaneity. *Physical Review Letters, 88*, 120404.

Stillfried, N. v. (2010). *Generalized entanglement: Theoretical and experimental explorations*. (PhD), Europa Universität Viadrina, Frankfurt (Oder).

Targ, R. (2019). What do we know about PSI? The first decade of remote-viewing research and operations at Stanford Research Institute. *Journal of Scientific Exploration, 33*, 569–592.

Targ, R., & Katra, J. E. (2000). Remote viewing in a group setting. *Journal of Scientific Exploration, 14*, 107–114.

van Fraassen, B. (2016). Naturalism in epistemology. In R. N. Williams & D. N. Robinson (Eds.), *Scientism: The new orthodoxy* (pp. 64–95). Bloomsbury.

Wackermann, J., Seiter, C., Keibel, H., & Walach, H. (2003). Correlations between brain electrical activities of two spatially separated human subjects. *Neuroscience Letters, 336*, 60–64.

Walach, H., Horan, M., Hinterberger, T., & von Ludacou, 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, 7*(2), 173–188. doi:10.1037/cns0000199

Walach, H., Lucadou, W. v., & Römer, H. (2014). Parapsychological phenomena as examples of generalized non-local correlations - A theoretical framework. *Journal of Scientific Exploration, 28*, 605–631.

Walach, H., & von Stillfried, N. (2011a). Generalised quantum theory—Basic idea and general intuition: A background story and overview. *Axiomathes, 21*, 185-209. doi:10.1007/s10516-010-9145-5

Walach, H., & von Stillfried, N. (2011b). Generalizing quantum theory - Approaches and applications. *Axiomathes, 21 (2)*(Special Issue), 185–371.

Walker, E. H. (1975). Foundations of paraphysical and parapsychological phenomena. In L. Oteri (Ed.), *Quantum physics and parapsychology* (pp. 1–53): Parapsychology Foundation.

Walker, E. H. (1979). The quantum theory of psi phenomena. *Psychoenergetic Systems, 3*, 259–299.

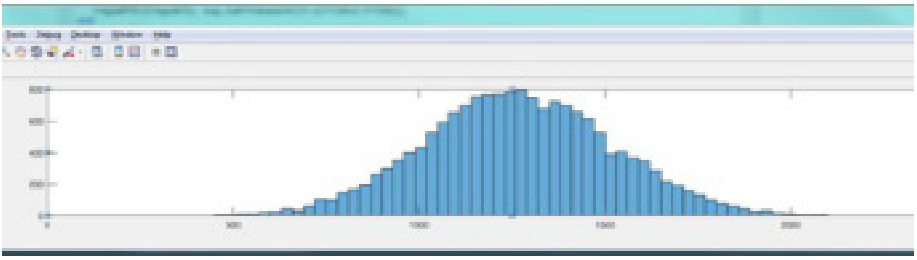
Walker, E. H. (2011, orig. 1974). Consciousness and quantum theory. In E. Mitchell (Ed.), *Psychic explorations: A challenge for science* (pp. 544–568). Cosimo.

Walleczek, J., & von Stillfried, N. (2019). False-positive effect in the Radin double-slit experiment on observer consciousness as determined with the advanced meta-experimental protocol. *Frontiers in Psychology, 10*(1891). doi:10.3389/fpsyg.2019.01891

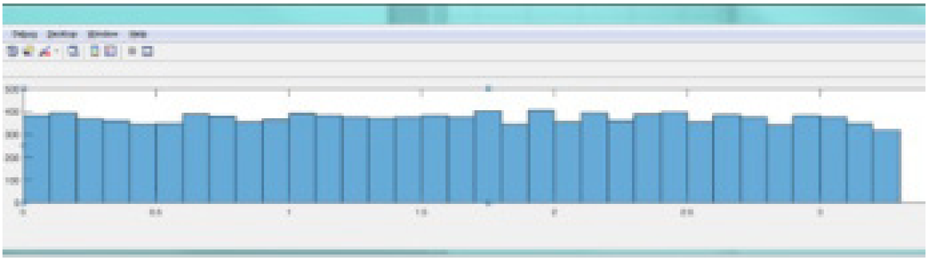
Watt, C., & Tierney, I. (2013). A preliminary test of the Model of Pragmatic Information using cases of spontaneous anomalous experience. *Journal of Consciousness Studies, 20*(11), 205–220.

Williams, R. N., & Robinson, D. N. (Eds.). (2016). *Scientism: The new orthodoxy*. Bloomsbury.

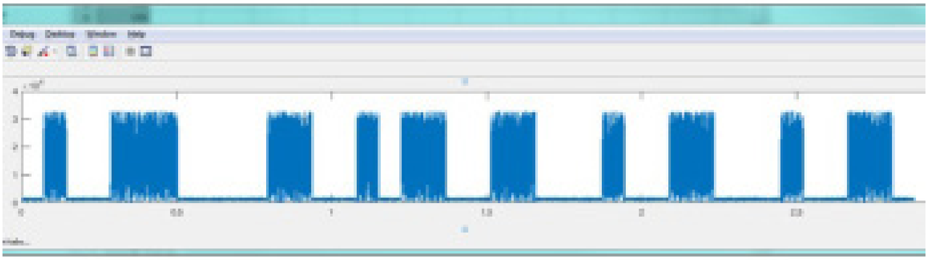
APPENDIX



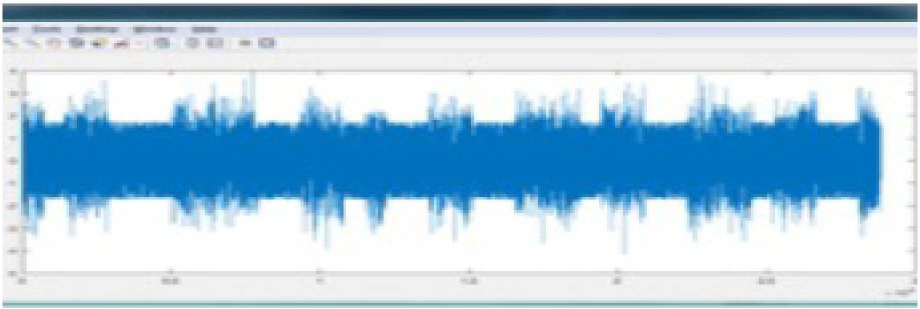
Appendix Figure 1. Distribution of sampling of True RNG.



Appendix Figure 2. Distribution of sampling of our traditional RNG.



Appendix Figure 3. REG-output of all REGs before normalization.



Appendix Figure 4. REG-output of all REGs after normalization.

APPENDIX TABLE 1

Number of Significant Matrix Elements in the 45 x 45 Experimental Matrix Compared to the Control Matrices C1 and C2 and to Chance Expectation Depending on Significance Level. Experiment 1 by KK, Original Analysis.

45 x 45 matrix	p-value (two-sided)	Number of significant correlations CE	Number of significant correlations C1	Number of significant correlations C2	Theoretically expected number of significant correlations	Difference between CE and C1	Difference between CE and C2
overall	0.1	246	226	241	203	20	5
part		115	83	121	99	32	-6
overall	0.05	141	111	134	101	30	7
part		52	32	70	50	21	-18
overall	0.01	39	22	19	20	16	19
part		11	1	7	10	10	4
overall	0.005	16	12	10	10	5	6
part		3	0	4	5	3	-1
overall	0.001	8	2	1	2	6	7
part		1	0	0	1	1	1
overall	0.0005	4	2	0	1	2	4
part		0	0	0	0	0	0
overall	0.0001	4	2	0	0	2	0
part		0	0	0	0	0	0

Because the data of this analysis were based on KK's own analytic strategy which is slightly different from that of TH who evaluated the data for this experiment statistically, some numbers deviate from Table 1.

APPENDIX TABLE 2

Statistical Analysis of Experiment 1 – 27*45 Matrix; Randomization Test with 10,000 Iterations

	27x45										
	sig_th	**0.1**	**0.05**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**	**0.0005**	**0.0002**	**0.0001**
full	**zo**	163.00	96.00	45.00	21.00	8.00	7.00	3.00	2.00	0.00	0.00
full	**n_sim**	697	439	511	1118	2482	585	990	676	1716	937
full	**p_sim**	0.0697	0.0439	0.0511	0.1118	0.2482	0.0585	0.099	0.0676	0.1716	0.0937
part	**zo_part**	83.00	36.00	13.00	8.00	1.00	1.00	0.00	0.00	0.00	0.00
part	**n_part_sim**	405	1817	2831	1903	5948	2727	3172	1879	838	446
part	**p_part_sim**	0.0405	0.1817	0.2831	0.1903	0.5948	0.2727	0.3172	0.1879	0.0838	0.0446

sig_th: theoretical significance level at which the number of significant correlations is counted
zo: number of significant correlations empirically found at respective level
n_sim: number of simulated matrices out of 10,000 with significant correlations at or above the number found empirically
p_sim: actual significance level of observed number of correlations (n_sim/10,000)
zo_part: number of correlations in time-forward (upper) part of the matrix
n_part_sim: number of significant correlations found in 10,000 simulations at respective level in upper part of the matrix
p_part_sim: actual significance level of observed number of correlations (n_part_sim/10,000) in upper part of the matrix

APPENDIX TABLE 3

Statistical Analysis of Experiment 1 – 18*27 Matrix; Randomization Test with 10,000 Iterations

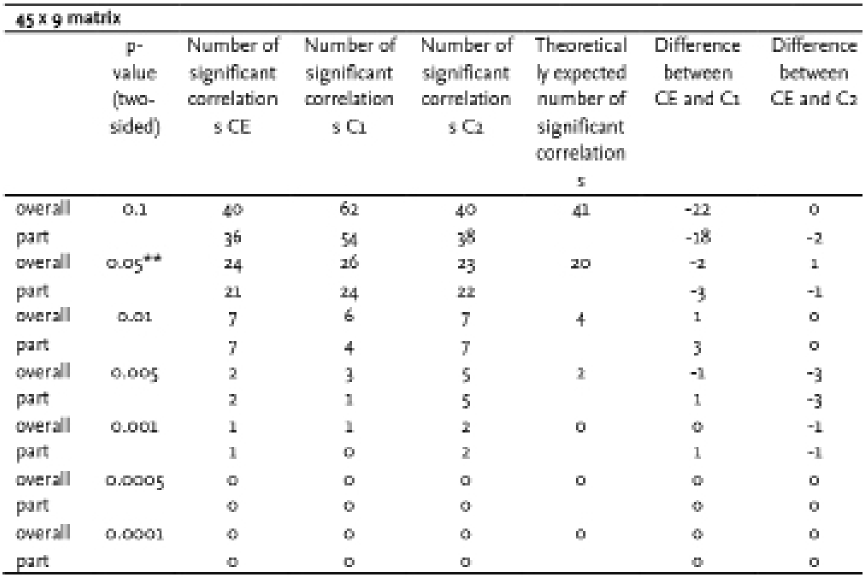
18x27

	sig_th	**0.1**	**0.05**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**	**0.0005**	**0.0002**	**0.0001**
full	**zo**	63.00	33.00	18.00	6.00	3.00	3.00	1.00	1.00	0.00	0.00
full	**n_sim**	1667	1876	1015	2631	2381	722	1071	490	788	427
full	**p_sim**	0.1667	0.1876	0.1015	0.2631	0.2381	0.0722	0.1071	0.049	0.0788	0.0427
part	**zo_part**	30.00	11.00	4.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
part	**n_part_sim**	1606	3888	3778	4961	5211	2732	1600	870	376	215
part	**p_part_sim**	0.1606	0.3888	0.3778	0.4961	0.5211	0.2732	0.16	0.087	0.0376	0.0215

sig_th: theoretical significance level at which the number of significant correlations is counted
zo: number of significant correlations empirically found at respective level
n_sim: number of simulated matrices out of 10,000 with significant correlations at or above the number found empirically p_sim: actual significance level of observed number of correlations (n_sim/10,000)
o_part: number of correlations in time-forward (upper) part of the matrix
n_part_sim: number of significant correlations found in 10,000 simulations at respective level in upper part of the matrix
p_part_sim: actual significance level of observed number of correlations (n_part_sim/10,000) in upper part of the matrix

APPENDIX TABLE 4

Number of Significant Matrix Elements in the 45 x 9 Varied Experimental Matrix with Psychological Variables Obtained by Questionnaire Compared to the Control Matrices C1 and C2 and to Chance Expectation Depending on Significance Level; Experiment 1 by KK, Original Analysis



45 x 9 matrix

	p-value (two-sided)	Number of significant correlations CE	Number of significant correlations C1	Number of significant correlations C2	Theoretically expected number of significant correlations	Difference between CE and C1	Difference between CE and C2
overall	0.1	40	62	40	41	-22	0
part		36	54	38		-18	-2
overall	0.05**	24	26	23	20	-2	1
part		21	24	22		-3	-1
overall	0.01	7	6	7	4	1	0
part		7	4	7		3	0
overall	0.005	2	3	5	2	-1	-3
part		2	1	5		1	-3
overall	0.001	1	1	2	0	0	-1
part		1	0	2		1	-1
overall	0.0005	0	0	0	0	0	0
part		0	0	0		0	0
overall	0.0001	0	0	0	0	0	0
part		0	0	0		0	0

Note. The number of correlations were calculated between 45 physical variables (TR, DT, KR, ZT, ZV x 9 runs) and 9 psychological variables (joy, love, anger, grief, fear, arousal, inner dialogue, direction of attention, absorption), reflecting the states of consciousness of the participants measured with the Phenomenology of Consciousness Inventory (PCI) (Pekala, 1995).

Program Code for the Permutation Test in Matlab:

```
for n = 1:10000
    % random permutations
    EPh2 = EPh(:,randperm(size(EPh,2)));
    CPh2 = CPh(:,randperm(size(CPh,2)));
    EPs2 = EPs(:,randperm(size(EPs,2)));
    CPs2 = CPs(:,randperm(size(CPs,2)));
    % calculation of correlation matrix
    [E_rho, E_p] = corr( EPh2’,EPs2’, ‘type’, ‘Spearman’, ‘rows’, ‘all’, ‘tail’,’both’);
    [C_rho, C_p] = corr( CPh2’,EPs2’, ‘type’, ‘Spearman’, ‘rows’, ‘all’, ‘tail’,’both’);
nc= size(E_p,1)*size(E_p,2);
ti=0;
sig_th = [.1, .05,.02,.01, .005, .002, .001, .0005, .0002, .0001];
n_soll = sig_th.*nc;
for p_th = sig_th
    ti=ti+1;
    psig=(E_p<p_th);
    no_exp(ti) = sum(sum(psig));
    psig=(C_p<p_th);
    no_cont(ti) = sum(sum(psig));
    zo(ti)=(no_exp(ti)-no_cont(ti))/sqrt(2*no_cont(ti)*(1-no_cont(ti)/nc));
    yo(ti)=(no_exp(ti)-no_cont(ti))/sqrt((no_exp(ti)*(1-no_exp(ti)/nc))+(no_
cont(ti)*(1-no_cont(ti)/nc)));
    do(ti)=no_exp(ti)-no_cont(ti);
    eo(ti)=no_exp(ti);
end
…
end
```