Journal of Scientific Exploration

Anomalistics and Frontier Science

RESEARCH ARTICLE

A Pre-Registered Test of a Correlational Micro-PK Effect: Efforts to Learn from a Failure to "Replicate"

Markus A. Maier*

Ludwig-Maximilians-Universität München, Germany
markus.maier@psy.lmu.de

Moritz C. Dechamps*

Ludwig-Maximilians-Universität München, Germany

*shared first authorship

SUBMITTED June 24, 2021
ACCEPTED February 22, 2022
PUBLISHED August 22, 2022

https://doi.org/10.31275/20222235

HIGHLIGHTS

We describe a failed replication of an erroneous finding in a previous psi / psychokinesis study. However, an advanced statistical analysis seemed to confirm a pattern we found in past studies. This result is consistent with the idea of unconscious experimenter psi.

ABSTRACT

Micro-psychokinesis (micro-PK) research studies the effects of observers' conscious or unconscious intentions on random outcomes derived from true random sources such as quantum random number generators (QRNGs). The micro-PK study presented here was originally planned, preregistered, and conducted to exactly replicate a correlational finding between two within-subject experimental conditions found in an original micro-PK dataset (n = 12,254) using a QRNG. However, after data collection and analyses, a data error was detected in the original to-be-replicated dataset. A reanalysis of the original correlation effect after error correction revealed strong evidence for the absence of a correlation in the original data. This study's primary goal was to test the existence of a correlational micro-PK effect in the present data as specified in the pre-registration. In addition to this replication attempt, the present study also can be considered an unsystematic case report or field study on experimenter psi (e-psi), since a strong expectation was formed initially about the occurrence of an effect that indeed was objectively absent from the original data. This study's results indicate no evidence for the existence of a correlational (and standard) micro-PK effect. In other words, the actual correlational data did not meet the experimenters' conscious expectations, and thus no consciously based effect of e-psi on micro-PK was found. However, the change in evidence for the effect across time described by the sequential Bayes factor resembled a data pattern that also was frequently reported by the experimenters in past studies. Although these data did not meet the criterion of statistical significance and a rejection of the null hypothesis failed, the marginal effects might be interpreted as weak influences based on unconscious e-psi. In addition, the trend observed matches both experimenters' general beliefs about the occurrence of e-psi in micro-PK. These findings' implications for the application of scientific methods to the study of micro-PK and psi in general are discussed.

KEYWORDS

Experimenter effects, experimenter psi (e-psi), micro-psychokinesis (micro-PK), quantum random number generator (qRNG)

INTRODUCTION

The present study was originally designed as a pre-registered experiment that aimed to replicate a statistical anomaly observed in the data of a series of three micro-psychokinesis (micro-PK) studies (Dechamps et al., 2021, Studies 1 to 3). These studies were conducted to test observer effects on quantum-based random number generators' (qRNG) outputs. Micro-PK research using qRNGs is a long-established tradition (Schmidt, 1970, 1974) and it explores human observers' abilities to mentally influence quantum-based outcomes. Numerous studies have been conducted with different variations in observers' intentions and various outcome measures, yielding an impressive amount of data. Two meta-analyses of these studies found an overall significant effect and thus evidence for observer-dependent variations in quantum randomness (Bösch et al., 2006; Radin & Nelson, 1989). Traditional micro-PK research to a greater extent applied active intention inductions by instructing participants to bias QRNG outcomes in a certain way (see, e.g., Jahn et al., 1997). A recent high-power micro-PK study using an active goal induction reported by Mossbridge and Radin (2021) found evidence for micro-PK in a first dataset that could also be replicated in a pre-registered analysis of similar data in a second dataset. These and other findings corroborate the usefulness of overt mental intentions in micro-PK.

In contrast to this approach, in the three precursor studies mentioned above, Dechamps et al. (2021) tested the impact of more implicit, unconscious processes on micro-PK outcomes using a subliminal priming technique in a within-subjects design. In the experimental condition during each trial, participants were subliminally primed with a pre-selected positive image before the qRNG chose between this same positive image or a negative counterpart. In the control condition, neutral priming was implemented. Each condition consisted of 20 trials involving the same 20 target pairs. All 40 trials were presented in randomized order. As predicted, in the first study (n = 4,092) the mean score of positive pictures based on qRNG selection exceeded chance expectation in the priming condition (*M* = 10.10; *SD* = 2.27), and Bayesian analysis revealed strong evidence for a micro-PK effect (BF_{10} = 13.35). Moreover, in accordance with their predictions, no decisive evidence for micro-PK (*M* = 10.03; *SD* = 2.26) was detected in the control condition (BF_{01} = 4.19). Later replication attempts (Studies 2 and 3) found no evidence for micro-PK in either condition (all BFs_{01} > 10; with one exception: in Study 2's control condition, the BF_{01} was > 4). The authors concluded that this data pattern obtained across the three studies could either be interpreted as a false positive in Study 1, owing to failed replications in Studies 2 and 3, or as a decline effect frequently observed in micro-PK research, corroborating the model of pragmatic information (MPI; von Lucadou, 1995, 2015, 2019; von Lucadou et al., 2007). In an additional exploratory analysis of the combined data from Studies 1 to 3 (total n = 12,254), the correlation between the mean scores of positive images obtained from the experimental and control conditions was calculated. This post hoc analysis revealed a correlation coefficient of *r*(12,254) = .032, BF_{10} = 36.46, indicating very strong evidence for a non-random correlation. Figure 1 depicts the sequential Bayes factors indicating the change of evidence of the micro-PK effects obtained in the experimental condition, in the control condition, and obtained from the correlation between both micro-PK effects.

After a visual inspection of the graphs, these authors hypothesized that the correlation effect would appear after evidence for micro-PK in the experimental condition had vanished. They concluded that as long as any standard micro-PK effects were absent, a stable correlation would be found. Thus, both authors had strong expectations that an exact replication would yield strong evidence for a correlation between micro-PK conditions. They set up a pre-registration in which they predicted the re-appearance of the above-described correlation in case a standard micro-PK effect was still absent from a newly to-be-collected dataset. Thus, the primary goal of the present study was to replicate the micro-PK findings reported by Dechamps et al. (2021, Studies 1 to 3) including a further test of the standard micro-PK effect found in the experimental condition of their original Study 1. Most central, however, the focus here was on the exact replication of the post-hoc finding described above: an a priori test of a positive correlation between the micro-PK data from both experimental conditions found across the dataset of Studies 1 to 3.

The authors conviction of the existence of such a correlational effect was based on their belief that strong Bayesian evidence has been found in the original dataset (Dechamps et al., 2021, Studies 1 to 3). This subjective expectation as documented in the pre-registration was therefore—in their view—supported by objective, although still preliminary data. The planned replication should further test and corroborate the objective nature of the correlational effect. However, the empirical background regarding the correlational effect and consequently the scientific substantiation of its appearance after closer inspection of the original data had to be revised. Before, during, and after the data collection and analyses of the study described herein, these authors did not know that the strong evidence for the correlation effect found in the data of Dechamps et al.'s (2021) Studies 1 to 3 was based on an erroneous coding of eight participants in their Study 3. The program coded the mean scores of positive images from subjects

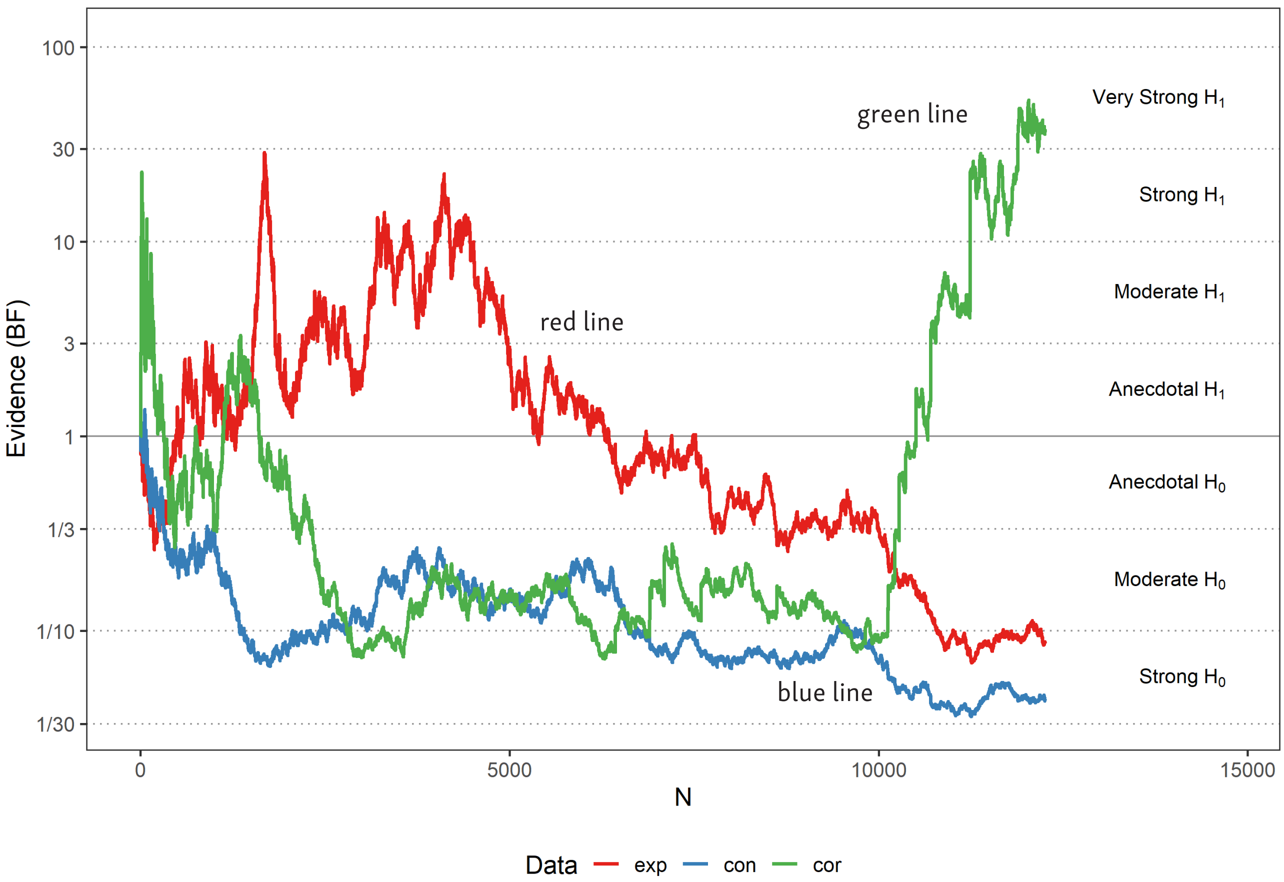

Figure 1. Erroneous sequential Bayes factors for Studies 1 to 3 from Dechamps et al. (2021) for experimental condition (red line), control condition (blue line), and correlation of both conditions' mean scores (green line).

who participated twice in a study with minus signs before the respective values. These data should have been deleted from all analyses but were mistakenly retained within the dataset. The authors detected this error after the present study's data had already been analyzed. After correcting the error and excluding the erroneous data, the original strong evidence for the correlation effect disappeared. As Figure 2 indicates, the real correlation effect in the data of Dechamps et al. (2021; Studies 1 to 3) was non-existent (r(12,254) = −.01), with a final BF_{01} = 49.48 indicating very strong evidence for a null effect. In other words, these authors' expectations and a priori hypothesis concerning the correlation effect in the present study were based on the erroneous assumption that a true effect existed. The propositions and predictions made in the pre-registration remained untouched by this new development, and the present data have been analyzed in accordance with all the specifications made in the pre-registration. However, on a theoretical level the interpretation of the present data in terms of replicability of an original effect turned into an unsystematic exploration of an experimenters' false expectations effect. Since these authors had a purely subjective expectation of the occurrence of an effect that stood in stark contrast to its true nonexistence, this new study only provided a case report or field study of potential experimenter effects in an actual micro-PK experiment. Although the unplanned nature of the investigation of the experimenter effect addressed here admittedly weakened the empirical value of this undertaking and cannot be viewed as an adequate scientific test of experimenter psi, the data can provide some preliminary and anecdotal insights into the question of which role purely subjective expectations of experimenters and data analysts might play in micro-PK (see, e.g., Rabeyron, 2020).

Experimenter effects have long been suspected of playing a significant role, particularly in micro-PK research (e.g., Kennedy & Taddonio, 1976). Varvoglis and Bancel (2015) argued that successful micro-PK investigators frequently showed outstanding micro-PK results when testing themselves. They also remarked that some experimenters' mindsets (e.g., Schmidt) exerted a biasing influence on micro-PK data obtained in their studies (see Varvoglis & Bancel, 2015, p. 278, footnote 8). Since micro-PK research using qRNGs explores observer effects on

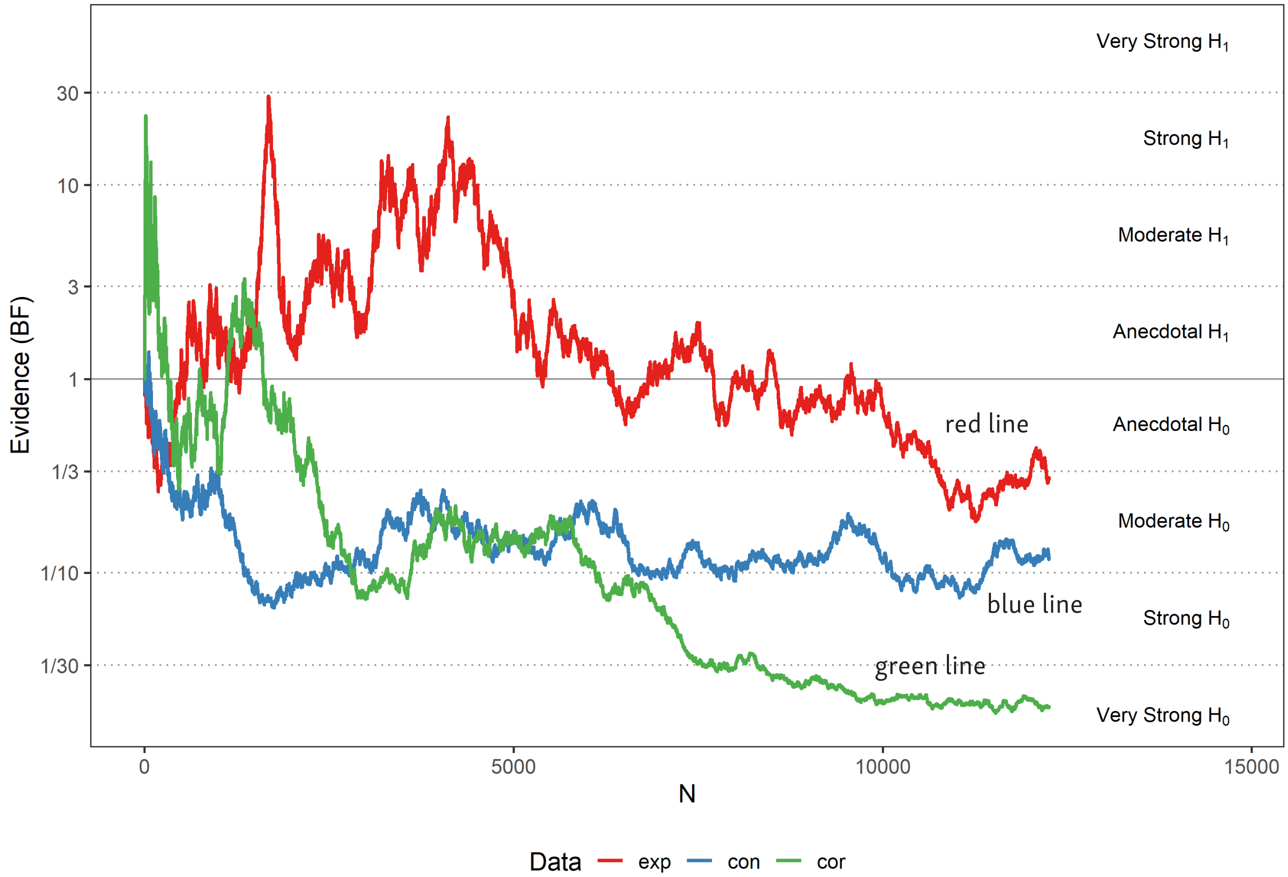

Figure 2. Actual sequential Bayes factors for Dechamps et al.'s (2021) Studies 1 to 3 for experimental condition (red line), control condition (blue line), and correlation of both conditions' mean scores (green line).

quantum-based outcomes, observation effects of this kind must not be restricted to the participants but should also be attributed to investigators and data analysts. This conjecture is central to the PMIR (psi-mediated instrumental response model (Stanford, 1977), which suggests that psi effects can arise unconsciously provided they fulfill an individual's (this includes experimenters) need. As Rabeyron (2020) has emphasized, this constitutes a major problem in research of this nature, as the scientific process typically demands that experimenters and their expectations be clearly distinguished from the phenomenon under review. In micro-PK specifically, it creates the problem that the true source of potentially successful results remains unknown. Although studies have reported anecdotal and indirect evidence of experimenter effects on micro-PK results (see Palmer, 2017), direct experimental tests of the occurrence of such effects in micro-PK remain rare, with the exception of a study conducted by Bierman (1978) that found no evidence for conscious experimenter effects in RNG data. The use of "silent" or "hidden" RNGs, which were concealed from the principal investigator during their studies, revealed significant results indicating experimenter psi (e-psi) (Berger, 1988; Honorton & Tremmel, 1979; Varvoglis, 1989; Varvoglis & McCarthy, 1986) and provided some preliminary evidence in this regard. However, additional research is required to further explore experimenter effects in micro-PK paradigms.

In sum, the present study's central goal was to test a standard micro-PK effect originally reported by Dechamps et al. (2021, Study 1) and to replicate a correlational micro-PK effect found with supposedly very strong evidence for H1 in a post-hoc analysis of the data of Dechamps et al. (2021, Studies 1 to 3). This effect consisted of a substantial correlation between the two within-subject conditions in a high-power micro-PK study involving subliminal priming. Moreover, no explicit reference to a decline effect was made. The study, including the original expectations and analytical methods, was preregistered at OSF (https://osf.io/a47g2), and data collection, reporting, and analyses followed the exact protocol outlined in this pre-registration. With regard to the empirical background concerning the original data and the interpretation of the present data, some adjustments had to be made. Since the original evidence for the correlational effect was based on faulty data and indeed this original dataset included a true null effect with respect to the correlation, the study unintentionally changed into an anecdotal field experiment unsystematically exploring experimenters' expectations on micro-PK correlations. The experimenters expected that strong evidence for the correlation ($BF_{10} > 10$) would be found under the condition that the standard micro-PK effect was absent in the experimental priming condition. Given the existence of experimenters' expectations that were indeed unknowingly in contradiction to the objective data, a relevant question concerning the role of experimenters' expectations in micro-PK could be anecdotally addressed and later discussed: Would such a false but strong belief itself affect the appearance of a correlation effect?

In additional analyses that were not part of the preregistration, we explored on a post hoc basis the stability of evidence for the correlation effect tested in this study. With additional data collection, evidence for this correlation will either remain stable or will decline after the initial evidence has been documented. In the latter case, *Change of Evidence* (CoE; see "additional analysis" section below) analyses will be applied to test the observed effect changes against random fluctuations in the evidence for the effect (false positives).

METHODS

Participants

Ethical Statement

Participants consented electronically to participation in the study by pressing an 'accept' button prior to the experiment. They were also informed in general terms about the study and advised that participation was voluntary and were given a brief written explanation of the study's purpose after its completion. All data were coded, stored, and analyzed anonymously. The procedure was approved by the ethical board of the Department of Psychology at LMU Munich and at Kantar (https://www.kantar.com/de), a data collection company specializing in online surveys that conducted the data collection.

Subject Pool

The study's sample comprised German participants distributed throughout the country. Participant recruitment and data collection were organized by Kantar. Kantar distributed invitations to participate in the study to a random selection of their participant pool daily via email, aiming for a completion rate of about 100 per day.

Statistical Approach and Data Collection

We adopted a Bayesian approach in this study. Bayesian inference statistics allow for data accumulation (i.e., the addition of individuals' data until a specific stopping criterion has been met). The correlation effect was assessed using a Bayesian correlation test. The following procedural details were specified in the preregistration: A BF of 10 indicating strong evidence for either H_0 or H_1 was defined as the stopping rule. In case the stopping criterion was not reached until approximately 1,000 participants had been tested, data collection would be stopped at this point. However, if a clear trend toward $BF_{10} = 10$ was recognizable at n = 1,000, we planned to increase this number of participants further.

The prior for the correlation analysis $\rho \sim$ Beta (0.1) was chosen a priori. This prior was based on an estimated effect size of r = .1 and has also previously been applied in the analysis of faulty correlation data (Dechamps et al., 2021).

Regarding standard micro-PK effects, we additionally tested whether the mean scores of positive pictures would exceed chance expectation in the experimental and control conditions using two separate one-sample Bayesian *t*-tests with a one-tailed approach for the analyses performed (applying an informed prior centered around 0.05 with an *r* of 0.05, i.e., $\delta \sim$ Cauchy [0.05, 0.05]). Although we a priori predicted and tested for the existence of a micro-PK effect, especially in the experimental condition, we were actually hoping for a null finding, as reported in Dechamps et al.'s (2021) Studies 2 and 3. The absence of a standard micro-PK effect in the data was considered to be a precondition for the occurrence of a correlation effect. This assumption was stated in the preregistration of this study.

The Bayesian analyses for the correlation and the standard micro-PK effects in the experimental and control conditions were performed on an irregular basis with the respective actualized sample's mean scores (more or less weekly within the first 1,000 participants and then again with the total sample of more than 2,000 participants). We used the statistical software R's 'Bayes Factor' package for the Bayesian analyses. The data collection took place between January 2021 and February 2021.

Sample Size

Although the stopping criterion of $BF_{10} = 10$ for the correlation effect was reached early on during data collection at n = 24, we considered the power of this subsample too small to provide convincing evidence for the effect. We thus continued data collection until the supplementary stopping criterion of n = 1,000 participants was reached. At this stage of data collection, the actual correlation was *r* = .034 and was within the range of the original result (*r* = .032). However, the corresponding Bayesian evidence remained inconclusive: BF_{10} = 0.42. We thus decided—in accordance with the preregistration—to collect additional data from a further 1,000 participants. We ceased data collection at this point, since the financial resources were exhausted. The total sample size was N = 2,052 (demographic data available for 1990 participants: 49.20% male, 50.55% female, 0.25% diverse; mean age = 44.22, SD_{age} = 13.84). Final Bayesian analyses were then performed with this complete set of data, as specified in the preregistration.

Materials

Experimental Program

The study was an exact replication of Dechamps et al.'s (2021) Studies 1 to 3. As in the original studies, the present study was run as an online experiment. All subjects could participate from any location using their private computers and internet access. The experiment was executed using a dedicated web server based in the university's computer center and displayed on the participants' web browsers. This was implemented using jsPsych (v 6.1.0; de Leeuw, 2015), a JavaScript library designed to run online behavioral experiments.

Stimuli

Positive and negative images were used as target stimuli and a mixture of them as prime stimuli. The target stimulus sets consisted of photographs obtained from Shutterstock, a provider of royalty-free stock images. The pictures used were pre-selected from this pool by the authors, both experts in experimental emotion induction techniques, using pictorial material. Stimulus selection was primarily based on independent valence estimations. Strongly negative and positive photographs were chosen based on the experts' ratings in case of mutual agreement. The positive target stimuli comprised 20 photographs depicting pets, peaceful landscapes, and groups of happy-looking people. Negative target stimuli encompassed 20 photographs depicting dangerous or attacking animals and other cataclysmal scenarios. The stimulus material was converted to black and white to balance out a general inequality with regard to the coloring of the positive and negative images. Both target sets were matched; that is, each positive target picture had a negative counterpart that was similar with respect to content. These pairs of target pictures represented specific subjects (e.g., a dog) with either positive (e.g., a friendly dog) or negative (e.g., an aggressive dog) valence.

From the target stimuli, two classes of priming stimuli were created. For the control (neutral priming) condition, each priming stimulus comprised an overlay of two matched target pictures. These primes were designed in such a way that the positive and negative stimuli were arranged with an equal emphasis (50/50). Therefore, the prime represented a homogenous mixture of both matched target pictures. Homogeneous mixtures of both target pictures were considered to constitute neutral primes since such arrangements were assumed to reflect the superposed existence of both affective states in the preconscious mind of the observer and would not activate any specific affective tendency above the other. Since 20 matched target pairs existed, the resulting number of corresponding priming stimuli was 20. Primes were accompanied by forward and backward masks comprising scrambled and indefinable versions of each prime. Each priming stimulus was presented three times during a given trial before the target display. The latter was randomly selected by a qRNG from the pair of targets from which the corresponding prime stimulus was created.

The experimental (positive priming) condition used the same mixtures from the matched target pairs and the same presentation modes during the trial, but following the first perfect 50/50 mixture presentation two slightly different priming images were displayed during each priming sequence within a given trial. In the second prime presentation, the positive share representing the positive target was displayed more distinctly than the negative share (60/40) and in the third prime presentation even more so (70/30). The positive image thus became more dominant during the priming sequence and was expected to be more strongly activated in the perceiver's unconscious mind. This rather unusual positive priming procedure should within a trial mimic the evolution of a classical reality and its conscious perception out of the preconscious mind of the observer under the biasing impact of an intentional goal (see Figure 3 for sample stimuli).

The assignment of experimental or control priming to a trial was performed using a pseudo-RNG (pRNG). Following the priming sequence, the quantum-based RNG (qRNG) randomly selected one of the two target images from which the priming stimuli were created in a given trial.

Generation of Quantum Randomness

During each trial after the priming sequence, a qRNG was used to determine whether the positive or negative image from the trial set was presented. To achieve this, a Quantis qRNG by ID Quantique was connected to the web server. This device generates two equally likely superposed quantum states by sending photons through a semi-conductive mirror-like prism. Upon measurement, only one of

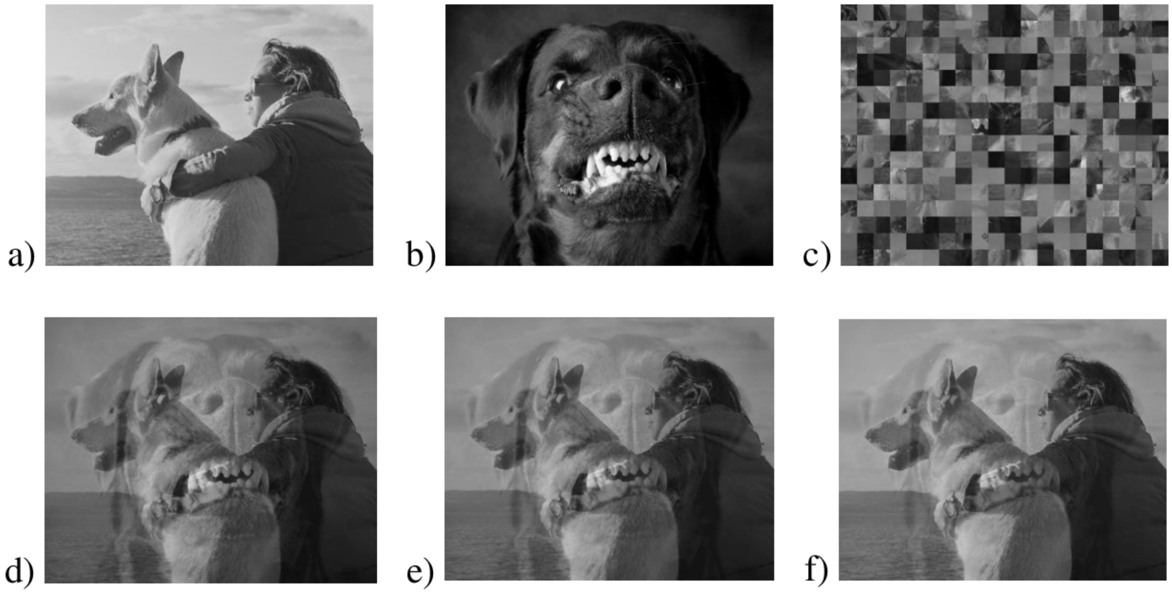

Figure 3. Sample stimuli. One trial consisted of a positive target picture (a) or a negative target picture (b), chosen by the qRNG. Scrambled-up versions of the 50/50 mixtures served as masking stimuli (c). Participants were primed with either an equal mixture of the two outcomes (d; neutral priming) or with a sequence of mixtures becoming more accentuated toward the positive target (d, e, and f; positive priming).

the two states can be observed and translated into either a 0 or a 1 bit. Using the random nature of quantum state reduction, a truly unpredictable result is generated. The qRNG consequently passed all major validation tests of randomness, such as the DIEHARD and NIST test batteries, and is regarded as one of the most effective sources of randomness (Turiel, 2007). The device was connected directly to the server via USB and generated a random bit for each trial after completion of the priming sequence and immediately before the display of the target stimulus, therefore working without a buffer. Care was also taken to ensure that each participant had received an individual bit.

Design

The study employed a within-subjects design with two conditions: an experimental condition (positive priming), in which the positive images from the respective matched target pairs served as the dominant share of the prime stimuli, and a control condition (neutral priming), in which neutral mixtures from respective matched target pairs served as prime stimuli.

Procedure

The invitation to participate in the study was issued via email by the polling company Kantar (https://www.kantar.com/north-america/about) to their pool of professional clients. Participants were advised to ensure an undisturbed environment before commencing the survey. They were asked for basic demographic information to ensure that they satisfied the inclusion criteria. They were then provided with a link that when clicked on took them to the experiment running on the university's web server. After the participants were asked to activate their browser's full-screen mode, they were shown written instructions for the task. Participants were advised that over the course of the experiment, they would repeatedly see flickering visual stimuli as well as positive and negative images and that these stimuli should be watched passively. They were reminded that they could abort the experiment at any time. Prime and image presentations began after the participants had acknowledged the instructions and had consented to participation.

Each participant viewed a total of 40 trials. For each individual, half of the 20 matched target pairs were randomly assigned to the experimental condition and the other half to the control condition using a software randomizer (pseudo-RNG) at the beginning of the experiment. Each of the 20 target pairs was used twice in this setting, resulting in a total of 40 trials. Next, the pseudo-RNG was used to individually permutate the order in which the 40 trials were presented via sampling without replacement.

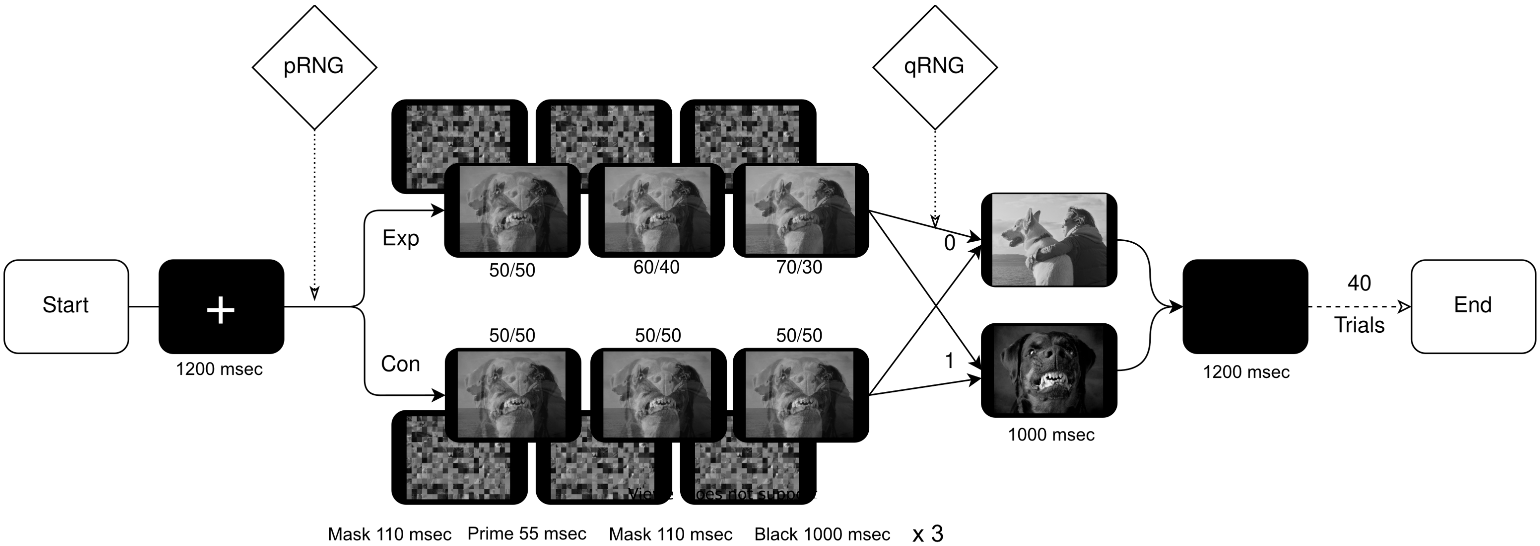

Figure 4. Schematic overview of the experimental design.

During each trial (see Figure 4), a fixation cross was first presented in the center of the screen (1200 ms) to direct the participants' attention toward this location. Next, in the priming sequence, a mixture (control condition) or various mixtures (experimental condition) of the two images that corresponded to the respective target pair of a given trial were used as prime stimuli. In the control condition, the 50/50 mixture prime stimulus was displayed three times for 55 ms each and in each instance was accompanied by a corresponding forward mask (110 ms) and a backward mask (110 ms) to ensure subliminal presentation. Each prime stimulus had a specific masking stimulus that was a scrambled version of the original. A 1000 ms-long gap showing a black screen was displayed between the three prime sequences. In the experimental condition, the presentation mode and times were the same as in the control condition, but the prime stimuli varied. The first prime presentation was a perfect mixture (50/50) of the target pair used in a given trial; the second prime presentation was a 60/40 mixture with the positive target picture being more visible; and the third prime presentation was a 70/30 mixture of this kind. In each trial, after the priming procedure was displayed, the qRNG was activated to provide an individual random bit that determined whether the positive or negative target stimulus from a given matched pair would be presented. The selected target picture was presented for 1000 ms. After this, a black inter-trial interval was presented for another 1200 ms before the next trial started. The two dependent variables consisted of the mean number of positive pictures and, therefore, the number of 0 bits generated by the qRNG in the experimental and control conditions.

RESULTS

First, evidence for the existence or non-existence of the standard micro-PK effects in both conditions were analyzed. Two separate Bayesian one-sample *t*-tests (one-tailed) were performed to test whether the mean number of positive images was higher than expected by chance in the experimental and control conditions. The expected mean score to occur by chance was 10 positive images (out of 20 possible) on average for each condition.

For the experimental condition, the Bayesian one-sample *t*-test (one-tailed) revealed a final BF_{01} = 11.00, indicating strong evidence in support of H_0. The mean score for positive pictures in this condition was M = 9.96 (SD = 2.26).

For the control condition, the Bayesian one-sample *t*-test (one-tailed) yielded a final BF_{01} = 13.95, indicating strong evidence in support of H_0. The mean score for positive images in this condition was M = 9.94 (SD = 2.26).

Figure 5 shows the sequential Bayesian analyses for both *t*-tests for each condition separately. Both sequential BFs showed a clear trend for a null effect, since the accumulated evidence increasingly supported the null hypothesis in both conditions.

Since a standard micro-PK effect was absent in both conditions, the assumed precondition required for a correlation effect to emerge was fulfilled. The Bayesian correlation (Bravais–Pearson) analysis between the mean scores of positive pictures obtained from the experimental and the control conditions yielded a correlation coefficient of $r(2{,}052) = .01$ with a final BF_{01} = 8.03, indicating moderate evidence in support of H_0. Figure 6 shows the sequential Bayesian analyses for the correlation.

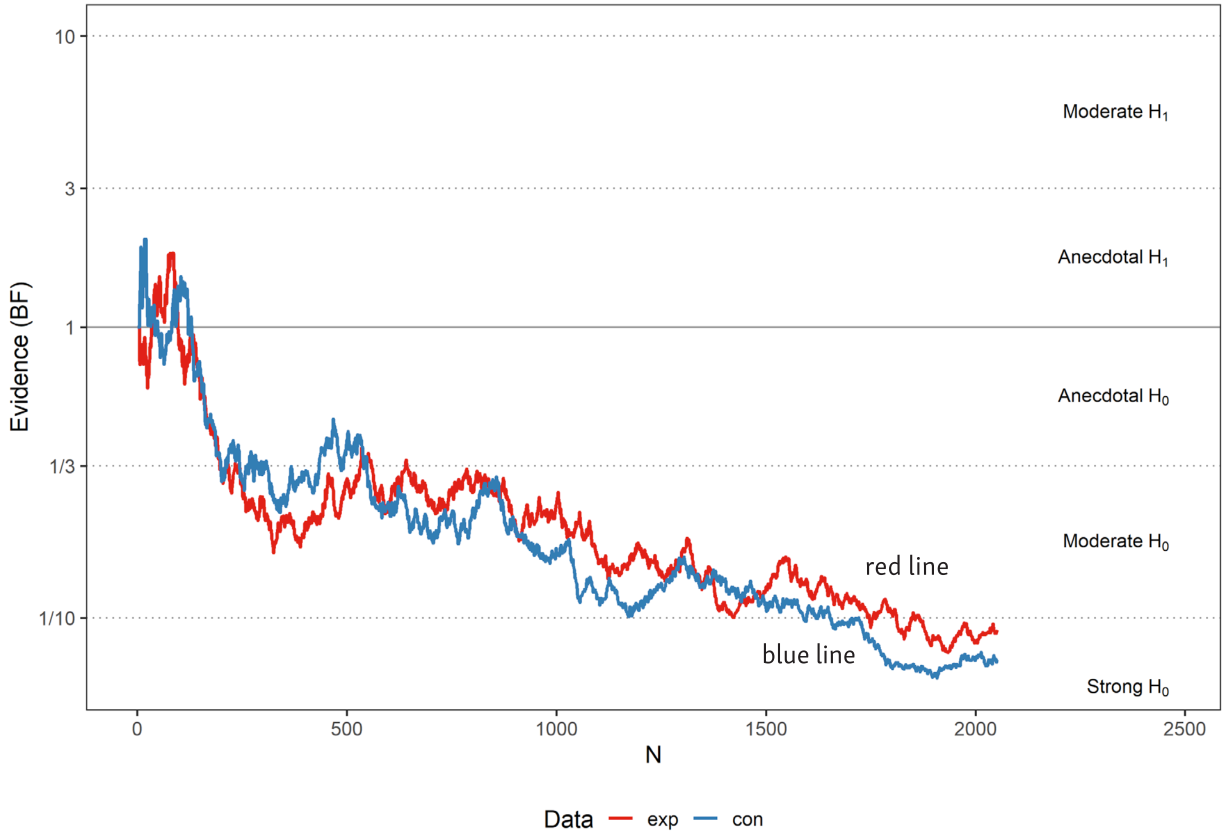

Figure 5. Sequential Bayes factors for experimental condition (red line) and control condition (blue line).

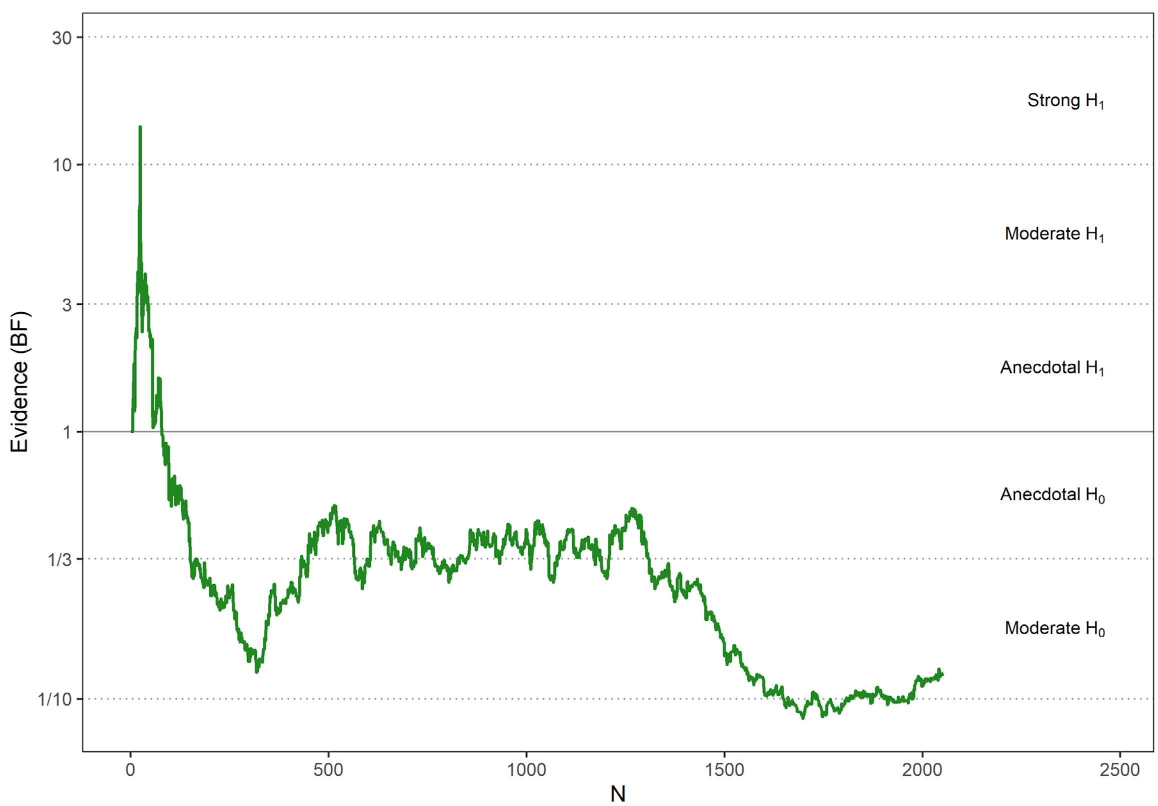

Figure 6. Sequential Bayesian correlation analysis of both conditions' mean scores.

As the graph illustrates, the stopping rule was met at n = 24 with a BF_{10} = 13.90 at this stage of the data collection process. In accordance with the preregistration, we could have ceased the data collection process at this point, on the grounds that strong evidence for the predicted effect had been found. However, we did not fully trust the results, in view of the sample size, which remained low at this phase of our investigation. We therefore increased the sample size to 1,000 participants. With this subsample, the BF still indicated moderate evidence in support of H_0, but the correlation coefficient of r = .034 appeared promising. A further 1,052 participants were recruited to closely monitor further changes in evidence for or against the effect. The correlation coefficient and its corresponding sequential BF declined further throughout this additional data collection process until the above-mentioned final BF was reached.

Additional Analyses

Although the above-mentioned stopping rule of BF_{10} > 10 was reached early on during data collection (at n = 24), these authors decided to carry on with data collection to test the robustness of the effect. At this time, they were still convinced that the effect remained stable. However, the results did not meet this expectation. The initial effect declined, finally indicating moderate evidence for a null effect.

Regarding the stability of the correlational effect across data collection, the presence of a decline in the effect was not explicitly mentioned in the preregistration and thus was not predicted. The authors tacitly assumed that this effect would hold true even if additional data were collected. However, two potential scenarios could be proposed: either strong evidence for this effect would remain stable until the data collection had concluded or, after strong evidence (BF_{10} > 10) had been reached, the effect might decline. The theoretical background for this second proposition is based on the MPI (von Lucadou, 1995, 2015, 2019; von Lucadou et al., 2007), according to which the initial indications of an anomalous effect are expected to decline during additional data collection. Such effect changes across time look identical to a data pattern that consists of an initial “false positive” later revealed as a “true negative.” To distinguish a “true positive + later decline” from such a “false positive + later true negative” dataset, Dechamps and Maier (2019) developed three tests, called “change of evidence (CoE) analyses”, performed on the sequential BF of the effect under investigation. These analytical techniques test the non-random nature of variations found in sequential BFs in comparison to sequential BFs obtained from simulated data (see Dechamps et al., 2021; Jakob et al., 2020). CoE analyses were not part of the preregistration but would serve as additional post hoc analyses of the correlation effect.

Thus, additional analyses were performed to test whether the change in evidence for the effect expressed by the sequential BF was based on non-random temporal variations. The CoE analyses consist of (a) an identification of the highest reach BF found within the experimental data at any time during the data collection compared with the highest BFs reached in 1,000 simulations of the data obtained from the same qRNG used in the original design (MaxBF analysis); (b) a test of the areas under the sequential BFs (energy of the curve), with BF = 1 as baseline obtained from the experimental data compared to the 1,000 simulations (BF energy analysis); and (c) a calculation of the sum of amplitudes of all frequencies obtained with fast Fourier transforms (FFTs) underlying the sequential BFs of the experimental data and the 1,000 simulations with a comparison of the sum of amplitudes obtained. These analyses test the non-random variation of the effect across time and provide a conservative test of non-random sequential BF fluctuations within such datasets.

The CoE analyses applied to the sequential BF of the correlation effect obtained in this study revealed the following results: the MaxBF analysis performed on the sequential BF of the correlation showed that the highest reached BF found in this dataset was BF_{max} = 13.90 at n = 24 and that 5.16% of the simulations had the same or a higher BF at any point within these datasets. Regarding the BF energy analysis, the sequential BF curve’s energy calculated as the area between the sequential BF curve and the BF = 1 horizontal line was –1309.49 for the correlation data, which was found to be surpassed by 26.13% of the simulations’ energies. The mean energy of the simulations was M = –784.82 (SD = 19604.62). For the FFT analysis, the sequential BF curves for the correlation effect and the 1,000 simulations were each Fourier-transformed with a sampling rate of 1/N. Since the transform is symmetric, only the first half is considered in the analysis, resulting in 1,026 tested frequencies for the correlation effect and for each simulation. To test the FFT results from the experimental data against chance occurrence, all 1,026 amplitudes obtained from the FFT of the experimental dataset were then added up, creating a sum score of the amplitudes obtained from all tested frequencies of this set. Similarly, the sum score of the amplitudes was computed for each of the 1,000 simulations. The amplitude sum of the experimental condition was 7.56, which was surpassed by 8.46% of the simulations’ amplitude sums. The mean amplitude sum of the simulations was M_{amp} = 7.40 (SD = 164.51).

In sum, none of the CoE analyses could reject the null hypothesis at the alpha = 5% level. Thus, the variations found in the sequential BF of the correlation data were not statistically different from random variations.

DISCUSSION

The study reported herein was originally designed to replicate a standard micro-PK effect reported by Dechamps et al. (2021, Study 1 to 3) and to replicate a correlation effect allegedly found between the mean scores of positive images in two within-subjects micro-PK conditions. Based on the apparently strong evidence in favor of the correlational effect (BF_{10} = 36.46), we set up a preregistration wherein we predicted that we would again find strong evidence (BF_{10} > 10) for such a positive correlation in cases wherein any standard micro-PK effects were absent (the latter was found to be true in the present data). We also supposed—but did not explicitly preregister—that this effect would remain stable across a sufficiently large data collection process. After the data collection was completed and the final Bayesian analyses of the data computed, we realized that the BF calculated from the to-be-replicated original dataset of Dechamps et al. (2021, Studies 1 to 3) was based on erroneous data. After the defective data points had been excluded, the correlation analysis yielded strong evidence for the null hypothesis (BF_{01} = 49.48). Consequently, the a priori prediction was based solely on the experimenters' expectations, and the present study involved a field experiment anecdotally exploring experimenters' expectations on micro-PK correlations.

Although a BF_{10} > 10 in line with the prediction for the correlation effect was found in the present data at an early stage of the data collection process, further data acquisition yielded moderate evidence in favor of the null hypothesis. This renders the interpretation plausible that no experimenter effects were present in the actual study. Specifically, the authors' stability assumption concerning the robustness of the effect was not satisfied. One could thus argue that strong and robust evidence for micro-PK effects cannot be produced solely by experimenters' expectancies (see also Bierman, 1978). The long-standing debate (see Bierman, 1978; Kennedy & Taddonio, 1976; Varvoglis & Bancel, 2015) as to whether and to what degree experimenter effects are significantly instrumental in micro-PK research at this point appears inclined toward the negative. This may be true, at least for the conscious expectations of Maier and Dechamps being present in that specific design and when focusing on the actual version of a micro-PK effect. It remains unclear whether this finding can be generalized to all experimenters, to unconscious expectations, and to the entire field of micro-PK research. In addition, the study presented here was not a priori designed as a test of e-psi. This lack of a systematic, scientific testing of experimenters' expectations reduces the informative value of the data in terms of providing evidence for or against e-psi to the level of an anecdotal report only.

With regard to unconscious expectations, the findings may be more ambiguous. Maier and Dechamps ran several micro-PK studies in the past, and in the majority of these experiments they found a data pattern of initial strong evidence and later decline of the effect (Dechamps et al., 2021; Dechamps & Maier, 2019; Jakob et al., 2020). We may thus envisage an implicit conviction on their part that micro-PK effects would usually follow this pattern. The CoE analyses of the sequential BF addressed this possibility. Descriptively, the sequential BF of the correlation effect exactly mimicked such a change of evidence in the data. Bayesian analyses of the actual correlation effect reached strong evidence for the effect (BF_{10} > 10) and then declined. However, none of the CoE analyses was significant—that is, the observed variation of the effect did not differ from random variations. This ruled out the possibility that unconscious experimenter expectation effects of this kind played a significant role in producing this data pattern. Nonetheless, a statistical trend was found in two of the three analyses, with one of the two tests being very close to the 5% level. Moreover, the strength of the correlation after the analysis of 1,000 participants closely resembled the expected effect size. This cast some doubt on the assertion that experimenter effects were truly absent in this study. Nevertheless, it may be safe to state that when the CoE results of this study are compared with those of past micro-PK studies conducted by our research team (Dechamps & Maier, 2019; Jakob et al., 2020), the extraordinary evidence of non-random changes in the sequential BFs of micro-PK data obtained in their research cannot be interpreted solely as e-psi but should be attributed to the participants' level of stimulus observation at least to a substantial degree. The experimenters' positive expectations might only additionally shape the micro-PK effect, making their research more successful (see also Parker & Millar, 2014) than if it were conducted by skeptics with negative expectations. Thus, facilitation or suppression effects might occur, depending on the respective expectation (Broughton, 1979), but these might not fully explain the observed effects. Rabeyron (2020) recently argued that e-psi imposes a practical limit on the scientific exploration of psi effects, making the determination of the effects' precise nature and the location of these effects' cause impossible (see also Broughton, 1979; Palmer, 1997). Regarding the conscious and unconscious expectations described above, our empirical data support this idea to some extent but also indicate that experimenters' expectations of this nature do not themselves produce psi effects—at least in our micro-PK research. On a side note, any evidence for micro-PK effects found regardless of the nature of their origin would nevertheless demonstrate that the mind creates reality.

Finally, the potential existence of a third form of the experimenter effect should be acknowledged. While writing up this discussion section, both authors retrospectively reflected on their overall attitudes toward e-psi. Their common belief was that e-psi effects simply played the marginal role of a moderating factor with small or negligible effect size; in other words, the data—as they are—were much appreciated by these authors. This attitude (or the *Journal of Scientific Exploration*'s readers' attitudes, etc.; see Rabeyron, 2020) may have produced the effects described herein. If this were the case, individuals' meta-level beliefs would have dramatic effects in psi research. The only possible escape from this dead-end scenario would be either to establish robust cynicism toward e-psi within the scientific community to reduce the potential impact of such effects or to accept that data obtained from this research (and perhaps from all scientific research) to some extent simply reflect (universal) beliefs in how the world should work.

At the end, some limitations of this study also have to be addressed. First, the unconscious processing of the subliminal prime presentations was not tested across participants with a standard test of subliminal perception such as a signal detection task. In contrast, the subliminal prime presentations' efficiency was only superficially explored by pre-testing the presentation mode in our research group and by our experiences with previous studies (Dechamps et al., 2021, Study 1 to 3). However, it was not considered crucial when some of the participants could identify the primes occasionally. In these potentially rare cases the positive picture was consciously pre-activated in the participant's mind and might have caused a similar tendency of approaching a positive target picture reality compared to a truly subliminal processing mode. Overall, we admit that a signal detection task performed after the micro-PK task to evaluate the unconscious processing of the primes would have provided a more stringent test. On the other hand, a pre-evaluation of the primes was considered sufficient for our goals and an inefficient subliminal priming procedure might not alternatively explain the null findings since this same procedure successfully revealed strong evidence for micro-PK in the past (Dechamps et al., 2021, Study 1). Secondly, the participants' compliance with the instructions and their attentional focus on the task were not controlled. Since this micro-PK task consisted of a purely passive watching task, inattentiveness during participation might have occurred and could have deteriorated the micro-PK effect. Thirdly, the prime and target pictures used in our design were chosen based on experts' ratings of valence only (see Dechamps et al., 2021) rather than using pictures from sets with normative ratings of valence and arousal. This suboptimal stimulus material could be another potential source for the null findings caused by confounding factors that counteract the valence effect.

In sum, our replication attempt of a standard micro-PK effect reported by Dechamps et al. (2021, Study 1) provided strong evidence for a null-effect in the present data. Also, in the present data and opposed to the present authors' prediction, moderate evidence for the null hypothesis was found with regard to a correlational micro-PK effect mirroring a similar null finding in the original data (Dechamps et al. (2021, Study 1 to 3). With regard to experimenter effects affecting the correlational micro-PK data, no or only minimal evidence for e-psi was found. The anecdotal nature of this aspect of our study prohibits further speculations regarding this question. Rather, future psi research should further explore the impact of e-psi on psi data in a more systematic way, within other areas of investigation and across different research teams (see Bierman & Jolij, 2020) to assess the extent of e-psi contributions in these investigative fields.

Resources

The study's preregistration, all digitally sharable material with the exception of the stimulus material, for which the authors do not hold the rights to distribute, and the experimental data and analyses scripts, are openly accessible at https://osf.io/tbha6.

REFERENCES

Berger, R. E. (1988). Psi effects without real-time feedback. *The Journal of Parapsychology, 52*(1), 1–27.

Bierman, D. J. (1978). Observer or experiment effect? A fake replication. *European Journal of Parapsychology*, *2*, 115–125.

Bierman, D. J., & Jolij, J. J. (2020). Dealing with the experimenter effect. *Journal of Scientific Exploration, 34*(4), 703–709.

Bösch, H., Steinkamp, F., & Boller, E. (2006). Examining psychokinesis: The interaction of human intention with random number generators—A meta-analysis. *Psychological Bulletin, 132*(4), 497–523. https://doi.org/10.1037/0033-2909.132.4.497

Broughton, R. S. (1979). Repeatability and experimenter effect: Are subjects really necessary? *Parapsychology Review, 10*, 11–14.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*, 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dechamps, M., & Maier, M. (2019). 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*(3), 406–434.

Dechamps, M. C., Maier, M. A., Pflitsch, M., & Duggan, M. (2021). Correlations between quantum states reductions and (un-)conscious states of observers: An empirical test and its replicability. Submitted to *Journal of Anomalous Experience and Cognition*.

Honorton, C., & Tremmel, L. (1979). Psi correlates of volition: A preliminary test of Eccles' "neurophysiological hypothesis" of mind-brain interaction. In W. G. Roll (Ed.), *Research in parapsychology 1978* (pp. 36–38). Scarecrow Press.

Jahn, R. G., Dunne, B. J., Nelson, R. G., Dobyns, Y. H., & Bradish, G. J. (1997). Correlations of random binary sequences with pre-stated operator intention: A review of a 12-year program. *Journal of Scientific Exploration*, *11*(3), 345–367.

Jakob, M.-J., Dechamps, M. C., & Maier, M. A. (2020). You attract what you are: The effect of unconscious needs on micro-psychokinesis. *Journal of Parapsychology, 84*(2), 227–253. https://doi.org/10.30891/jopar.2020.02.06

JASP Team. (2018). *JASP (Version 0.9) [Computer software].*

Kennedy, J. E., & Taddonio, J. L. (1976). Experimenter effects in parapsychological research. *Journal of Parapsychology, 40*(1), 1–33.

Mossbridge, J. & Radin, D. (2021). Psi performance as a function of demographic and personality factors in smartphone-based tests: Using a "SEARCH" approach. *Journal of Anomalous Experience and Cognition*, *1*(1–2). https://doi.org/10.31156/jaex.23419

Palmer, J. (1997). The challenge of experimenter psi. *European Journal of Parapsychology, 13*, 110–125.

Palmer, J. (2017). 'Experimenter Effects'. *Psi Encyclopedia*. The Society for Psychical Research. https://psi-encyclopedia.spr.ac.uk/articles/experimenter-effects

Parker, A., & Millar, B. (2014). Revealing psi secrets: Successful experimenters seem to succeed by using their own psi. *Journal of Parapsychology, 78*(1), 39–55.

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). https://doi.org/10.3389/fpsyg.2020.562992

Radin, D. I., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics, 19*(12), 1499–1514. https://doi.org/10.1007/BF00732509

Schmidt, H. (1970). A PK test with electronic equipment. *Parapsychology, 34*, 175–181.

Schmidt, H. (1974). Comparison of PK action on two different random number generators. *Journal of Parapsychology, 38*(1), 47–55.

Stanford, R. A. (1977). Conceptual frameworks of contemporary psi research. In B.B. Wolman (Ed.), *Handbook of parapsychology* (pp. 823–858). Van Nostrand Reinhold.

Turiel, T. P. (2007). Quantum random bit generators. *The American Statistician, 61*(3), 255–259.

Varvoglis, M. (1989). A "psychic contest" using a computer-RNG task in a non-laboratory setting. In L. A. Henkel & R. E. Berger (Eds.), *Research in parapsychology 1997*. Scarecrow.

Varvoglis, M., & Bancel, P. A. (2015). Micro-psychokinesis. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century.* (pp. 266–281). McFarland.

Varvoglis, M. P., & McCarthy, D. (1986). Conscious-purposive focus and PK: RNG activity in relation to awareness, task-orientation, and feedback. *Journal of the American Society for Psychical Research*, *80*(1), 1–29.

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

von Lucadou, W. (2015). The correlation matrix method (CMM): A new light upon the repeatability problem in parapsychology. Paper for the 58th Annual Convention of the Parapsychological Association and the 39th SPR International Conference, University of Greenwich, London.

von Lucadou, W. (2019). Homeopathy and the action of meaning: A theoretical approach. *Journal of Scientific Exploration, 33*(2), 213–254.

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