Testing the potential paradoxes in “retrocausal” phenomena

Jacob Jolij, and Dick J. Bierman

Citation: AIP Conference Proceedings **1841**, 030003 (2017); doi: 10.1063/1.4982774
View online: https://doi.org/10.1063/1.4982774
View Table of Contents: http://aip.scitation.org/toc/apc/1841/1
Published by the American Institute of Physics

Testing the Potential Paradoxes in "Retrocausal" Phenomena

Jacob Jolij [1, a)] and Dick J. Bierman[1, b)]

[1] *University of Groningen, Heymans Building, 9712 TS Groningen, The Netherlands*
[b)]*Corresponding author: d.j.bierman@rug.nl*
[a)]*j.jolij@rug.nl*

Abstract. Discussions with regard to potential paradoxes arising from "retrocausal" phenomena have been purely theoretical because so far no empirical effects had been established that allowed for empirical exploration of these potential paradoxes.
In this article we describe three human experiments that showed clear "retrocausal" effects. In these neuropsychological, so-called, face-detection experiments, consisting of hundreds of trials per participant, we use brain signals to predict an upcoming random stimulus. The binary random decision, corresponding to showing a noisy cartoon face or showing only noise on a display with equal probability is taken after the brain signals have been measured. The prediction accuracy ranges from 50.5-56.5% for the 3 experiments where chance performance would be 50%.
The prediction algorithm is based on a template constructed out of all the pre-stimulus brain signals obtained in other trials of that particular participant. This approach thus controls for individual difference in brain functioning.
Subsequently we describe an experiment based upon these findings where the predictive information is used in part of the trials to determine the stimulus rather than randomly select that stimulus. In those trials we analyze what the brain signals tell us what the future stimulus would be and then we reverse the actual future that is presented on the display. This is a 'bilking' condition. We analyze what the consequence of the introduction of this bilking condition is on the accuracy of the remaining (normal) trials and, following a suggestion inferred from Thorne et al, we also check what the effect is on the random decision to either bilk or not bilk the specific trial. The bilking experiment is in progress and the results so far do not allow for conclusions and are presented only as an illustration.

INTRODUCTION

In recent years several authors reported anomalous correlations between physiological measurements during a pre-stimulus period and the random conditions of the subsequent stimulus [1]. These findings suggest either experimental errors and the use of 'questionable research practices' or a "retrocausal" effect. The phenomenon has been reported for many different physiological variables, most notably with dermal activity variables. But also variables like BOLD, ECG and pupil dilation have been shown to correlate with random futures.
In the philosophy of science such a series of observations, apparently not fitting in the currently accepted world view, has been referred to as 'an anomaly'. This particular anomaly has been called 'presentiment' in the parapsychological literature.
Retro-causality in the physical world is an increasingly discussed topic. Ever since the end of the previous century publications on "retrocausality" have exponentially exploded (see fig. 1). In Newtonian physics there seems to be no place for "retrocausality but in the theory of relativity time travel into the past is not excluded while in almost all modern descriptions of the" physical world, most notably in the theory of electromagnetism (EM) time-symmetry is clear from the mathematical descriptions. Time symmetry is however not identical to "retrocausality" (see for instance Ruth Kastner, in this volume).

Quantum Retrocausation III
AIP Conf. Proc. 1841, 030003-1–030003-12; doi: 10.1063/1.4982774
Published by AIP Publishing. 978-0-7354-1511-9/$30.00

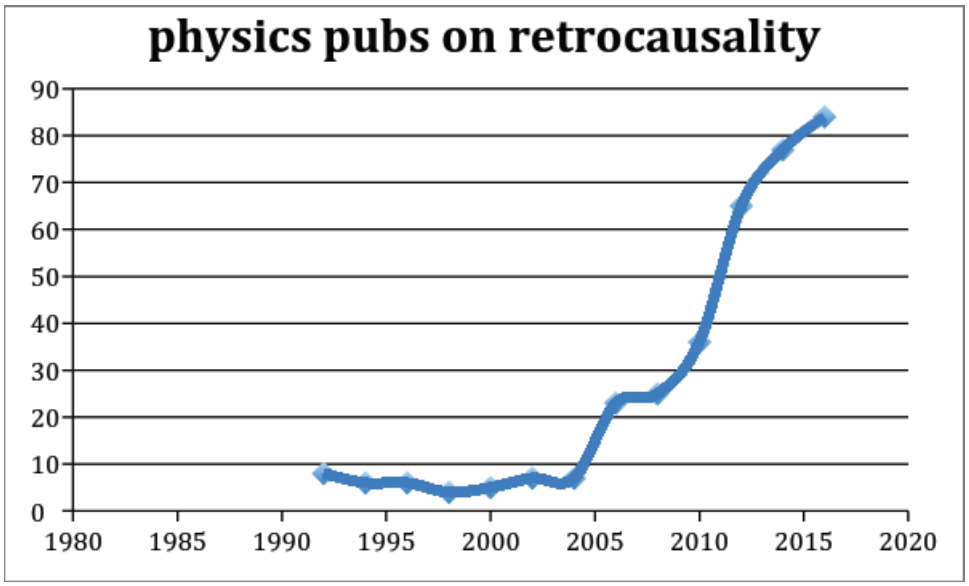

FIGURE 1. Number of publications per 2 year on "retrocausality" in physics journals

According to Popper [2], science can make major steps once an anomaly is established. However the status of the database of this type of presentiment experiments is far from 'well-established'. The beauty of this particular anomaly is that it should be measurable in almost every psychophysiological experiment with conditions that are truly randomly assigned. In this paper we will first re-analyze three straightforward EEG studies on face detection that were done during the last years at the Heymans lab of the University of Groningen. In the 'presentiment' literature there are 6 studies that have used EEG measurements [3-8]. Most of these have not been incorporated in the 2012 meta-analysis because that analysis excluded studies that did not report a significant response (i.e. normal) effect. The rationale for exclusion can be found in the review paper and is shortly discussed in the section 'Is there time-symmetry' in this paper.

Reviewing these EEG studies is near impossible because in each of these studies there was freedom to select EEG-channels or select sub-groups from the tested population. From the reports it is unclear if this occurred post hoc. Almost all studies reported SOME effect. But we cannot combine these studies because of the mentioned danger of over-analysis but also because of large differences between the studies. In some studies the subject was totally passive while in others the subject had an active role. So these results should be taken with a grain of salt because they may be due to publication bias and other questionable research practices (QRPs). In a recent simulation of another paradigm in the parapsychological literature it was found that, under reasonable assumptions with regard to the incidence of QRPs, about 60% of the overall effect could be explained away [9]. Therefore it is mandatory to eventually repeat these experiments that are supposed to be suggestive of an anomaly in the context of a pre-registered multi lab replication project with measures and standardized procedures that would exclude any questionable research practice. Over-analysis can be excluded by first analyzing a part of the trials and use the effects therein as a predictor for the other trials.

The three face detection studies in this paper all use a procedure where for each subject a part of the trials is used to predict the effect in the remaining trials using a processing method that does not allow for any post hoc selection by the experimenter. These three are all the studies on this topic done at our department and no subjects were removed post hoc. The number of subjects was not precisely fixed before the study but was set by other criteria unrelated to the results. The studies were not pre-registered and should be considered to be explorative.

We wondered if these studies would also show a presentiment effect and if so what the effect size would be. If a reasonable effect size could be established then it would in principle be possible to use this approach to do empirical research on the potential paradoxes that may arise through "retrocausal" effects (or time travel). In the second part of this paper we will sketch such an empirical approach and present the very first results.

A major methodological question is if the pre-stimulus EEG patterns that are assumed to predictive for the future stimulus vary per subject or if there is a general pattern across subjects. In almost all studies to date the latter has been assumed. Physiological processes can however be idiosyncratic in the sense that a majority of participants may respond with for example an increase in physiological activity in a specific context while a minority does the opposite. Averaging results over all participants then tends to obscure any effect. In the experimental approach we assume that subject do differ with respect to their brain signal patterns and our analyses take these differences into account.

The remainder of this article consists of two major parts. In the first part we describe the EEG face detection experiments in detail with the results that strongly suggest a "retrocausal" component. In the second part we show how by slightly changing the set-up of the experiments, so that the brain signals can be classified in real time, we have implemented a bilking experiment that aims at exploring the effects of deliberately creating a time-loop paradox in this experiment.

PART I: THE FACE DETECTION EXPERIMENTS

Methods

General

We present data from three different studies that each used the same visual stimulation protocol, but differed in procedure and general context. The visual stimulation protocol is identical to the one used in [10,11] and described in detail below. Data for the three studies has not been published elsewhere yet, but is aimed at analyzing stimulus processing related activity. In this paper, we report anomalous EEG activity during the pre-stimulus period in each of these three experiments that seems to be related to anticipation of future random conditions.

The three studies have been carried out between 2014 and 2016 at the University of Groningen, in the lab of the first author of this paper. The first study was about the effects of mood and caffeine consumption of visual perception, and will be called the 'Coffee' study from hereon; the second study dealt with the effects of hypnosis on visual perception, from hereon the 'Hypnosis' study'; the final study was about the effects of social influence on visual perception in romantically engaged couples, from hereon the 'Couples' study.

Participants

Table 1 gives the particulars of the participants in each of the three studies.

A total of 102 students (50 women, 52 male, mean age 21.5 years), generally recruited from the first year psychology program of the University of Groningen, participated in these 3 studies. All participants reported normal or corrected-to-normal vision, no history of physical or mental health problems, and gave written informed consent to participate. These studies were approved by the local Ethics Committee ("Ethische Commissie van het Heymans Instituut voor Psychologisch Onderzoek") under codes 13057-NE (Coffee), ppo-013-042 (Hypnosis), and ppo-015-053 (Couples)

Study	Year	N-blocks	N-male	N-female	Population	mean age	Sd age
Hypnosis	2015	8	12	15	Freshman students	20.4	1.6
Couples	2016	8	27	27	Couples (students)	22	2.6
Coffee	2014	6	11	10	Freshman students	21	1.6
total	-	-	50	52	-	21.5	2.2

TABLE 1. A review of the face detection studies. In the coffee-study 26 EEG channels were used whereas in the hypnosis and couples study only 8 channels are used. One block equals to 120 trials.

Stimuli and apparatus

The visual stimuli were identical in all three studies, and were presented using a Windows 7 PC with a 22" Iiyama TFT screen, running Matlab R2011a (The Mathworks, Natick, USA) and Psychophysics Toolbox version 3 (Brainard, 2003). Stimuli were schematic faces embedded in dynamic noise. Every trial was a short clip of 9 frames, each frame presented for 100 ms. Every frame contained an array of white noise, resulting in an animation of moving noise. The middle frame could either contain a schematic face or be left blank, and was accompanied by an rectangle around the array of noise serving as a cue that the critical stimulus could appear in this frame (See fig. 2). The time from the end of a trial to the beginning of the next one is jittered between 3.5 and 6.5 seconds, including the response time of the participant. The response time-out is 3 seconds. Between the blocks subjects were requested to rest for about 4 minutes (3 minutes minimum).

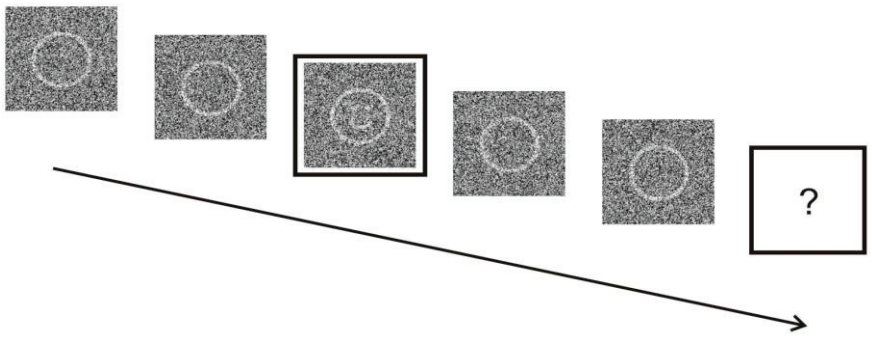

FIGURE 2. Schematic example of one face detection trial over time. Rather than showing all noise frames (4 before and 4 after the critical frame) only 2 frames, before and after, are shown. The cue (rectangle around the frame) indicates the critical frame. The question mark indicates the subject has to respond 'face seen' or 'no-face seen'. Time between stimuli is ~5 seconds.

Per trial, presentation of a target was triggered by drawing a random number from either the built-in pseudo random number generator of Matlab, based on the Mersenne Twister ('Coffee' study), by drawing a random number from the TrueRNG2 (ubuild.it, USA), a hardware number generator based on quantum avalanche noise in a diode circuit ('Hypnosis study'), or by means of a Geiger counter-based RNG ('Couples' study). When the drawn number

was even, a blank trial was presented, when the number was odd a face trial was presented, resulting in an overall target probability of 50%.

EEG data Collection and Processing

In the Hypnosis and Couples studies, EEG data was collected using a Porti 8 channel system (TMSi, Enschede, The Netherlands). Electrodes were placed at 10/20 locations T7, T8, P7, P8, Pz, O1, O2, and Oz. Data was sampled directly into Matlab with a sample-rate of 250 Hz and stored to disk on a trial-by-trial basis. Stimulus onset was marked using a digital trigger; to avoid possible 'contamination' of the EEG signals before stimulus onset the same trigger value was used for 'face' and 'blank' trials. For the Coffee study, data was collected using a TMSi ReFa 32 channel EEG system (TMSi, Enschede, The Netherlands). Electrodes were placed at the standard 10/20 locations, plus Oz. Data was recorded using Portilab version 2.0 (TMSi, Enschede, The Netherlands), and analyzed in Matlab.

Trials were filtered between .01 and 40 Hz using a zero-phase digital second-order causal Butterworth filter, baseline corrected (baseline from 800 ms to 600 ms pre-stimulus), and DC-detrended. To ensure that there were no artefacts produced by this filter we checked the temporal distortion by the filter. This was done by filtering a block signal; the filtered signal did show a minor pre-stimulus distortion from t = - 40 to 0 ms; therefore we may assume the prestimulus classification window that spans the interval from -600 till -100 ms was not affected by any filter artifacts. Trials with a peak amplitude >35 uV were excluded from analysis. Including these trials does not affect the results, though.

Binary data classification was done using a nearest-mean classifier using all available channels [12-14]. This method is well suited for the single trial detection of evoked potential components, such as the P300 or the N170 [12]. Evoked potential paradigms as used in the studies presented here are typically subtraction paradigms, in which the difference in brain activity between two conditions is of interest. In the present paradigm, we have blank and face trials, which only differ with respect to the presence of a target. Of interest is the additional activity evoked by the face trials as compared to the blank trials. Mathematically, we may express the relation between these evoked waveforms as in equations 1 and 2:

$$1.\, f(t) = b(t) + a(t)$$
$$2.\, a(t) = f(t) - b(t)$$

With f(t) being the activity evoked by a face presentation, b(t) the activity evoked by a blank presentation, and a(t) being the additional activity related to face processing. In order to classify an individual trial x(t) as a face or a blank trial we need to establish whether this trace contains a(t). Given that the sign of a(t) can vary over time, and contains important information, an efficient way of computing a classification value for the interval t_{min}, t_{max} is to compute the product of x(t) and a(t), and take the integral of this product (equation 3):

$$3.\, S\big(x(t)\big) = \int_{tmax}^{tmin} x(t)a(t)$$

This method will result in positive values when x(t) contains a(t), and in values close to or below zero when x(t) does not contain a(t). We can now make a binary classification by comparing S to a constant threshold. To find the optimal threshold, we first compute S(f(t)) and S(b(t)), giving us the extreme values S can take. The optimal threshold to discriminate between is then

$$4.\, C = \frac{S\big(f(t)\big) + S(b(t))}{2}$$

Classification outcome is then

$$5.\, c(x(t)) \begin{cases} 0 \; if \; S(x(t)) < C \\ 1 \; if \; S(x(t)) \geq C \end{cases}$$

Classification templates were constructed using a 'leave-one-out' approach, i.e., for every trial a new classification template was computed, excluding the trial to be classified. This way, a trial is never included in the template used to classify that trial.

For post-stimulus classification, we used the window between 100 ms and 600 ms after stimulus presentation; pre-stimulus period classification was done on the window from 600 ms to 100 ms before stimulus presentation.

Procedure

Participants were welcomed into the laboratory, and were briefed on the task.

In the hypnosis study they started with a brief practice session of 10 trials during which no data was recorded. After this, they did two blocks of 120 trials as a control measurement; subsequently they did two blocks after listening to a prerecorded suggestion of 'blurred vision' (e.g., they were given the suggestion "no faces are to be seen", even though in 50% of trials faces were presented); two blocks after listening to a prerecorded suggestion of 'clear vision' (e.g., they were given the suggestion "every trial contains a face", in reality this was 50%); the session was concluded with two blocks as a post-measurement. The order of suggestions was counterbalanced between subjects.

In the Coffee study, participants were tested on two separate sessions; on one of the sessions they were given decaffeinated coffee, on the other session regular coffee. 45 minutes after coffee administration, they did six blocks of the task. Two blocks were recorded after a positive mood induction; two blocks were recorded after a negative mood induction, and two blocks were recorded in a neutral mood. Because mood is known to affect expectancy processes in this task (see [10]) we only analyzed the neutral block (i.e., two blocks).

Finally, in the couples study, participants came in pairs. Each individual was seated behind their own computer, but would see the same stimuli. However, first participant A would see the stimulus, and then respond, subsequently participant B would see A's response as cue, and then the stimulus. After one block, the order would reverse, and B would go first, etc. For the results presented here, we of course only analyze the uncued trials (i.e., trials in which the participant would be first to respond). In total, participants did eight blocks of trials.

Data Analysis

After single trial classification of the trial using the prestimulus period template, we computed the proportion correctly classified trials per subject (i.e. face trials classified as face, and blank trials classified as blank), both for the prestimulus and for the post-stimulus intervals. As a control upon this crucial part of the analysis, raw data were provided to an independent analyzer who using his own software confirmed the results of the analysis. Proportions were tested against chance level (mu = .50) using a single value t-test to assess classifier accuracy. Moreover, we computed Pearson correlation coefficients between pre- and post stimulus classifier accuracy.

Results

Predictive Accuracy

The results of the three studies are presented as Accuracy values (Table 2). The accuracy value is the percentage of correctly predicted future conditions. Chance level is at 50%.

Study-name	N-Ss	Accuracy		t-value (prediction accuracy compared with 50%)	Effect-size (r)
		pre	post		
Hypnosis	27	50.46	61.44	1.02 (df=26)	0.20
Couples	54	52.35	60.16	5.08 (df=53) **	0.57
Coffee	21	56.27	73.63	5.43 (df=20) **	0.77
TOTALS	102	52.66	63.27	6.55 (df=101)**	0.64

TABLE 2. Accuracy of predictions based upon the pre-stimulus derived classifier and the post-stimulus derived classifier. T-values for a study are calculated using the prestimulus classification accuracies per subject with MCE = 50%. ** = p << 0.0001

To control for classifier bias, we simulated 100 experiments containing only noise trials (240 trials per experiment in 26 channels), and randomly labeled these trials as 'target' or 'no target' trials. Subsequently, we ran our classification procedure on this data. This resulted in an average accuracy of 49.32% over all 250 simulated experiments [t = -1.8877, df = 99, p = 0.0620]. Therefore, if there is any bias in the classifier, it is too conservative rather than too liberal.

When using the trial number as a factor in the analyses there are no interactions with that factor and hence there are no significant inclines in accuracy within a subject over the trials, suggesting that no learning occurred.

Theoretical Considerations

The reported effects do support earlier findings in the presentiment literature. Most of the effects reported in that literature use as random future conditions an emotional versus a neutral one. In a study on the effect of future feedback on the duration of one of two (spontaneously switching) percepts of Necker cube, Bierman found however that emotion apparently needn't to be involved in retrocausality [15].

In models that use arguments about (broken) time-symmetry to account for presentiment results like in CIRTS [16] one wouldn't expect that this kind of 'retrocausal' effect are limited to futures where emotion is one of the alternatives. However one would expect that if the EEG responses differ strongly in the face and the no-face condition this would also be the case for the presponses (an alternative name for the activity during the pre-stimulus period). We therefore calculated the correlation between the pre-stimulus derived and the post-stimulus derived (response) accuracies. The results are given in table 3.

TABLE 3. Correlations between accuracies based upon the pre-stimulus derived templates (psi_A) and the accuracy based upon the response templates (response_A). * = p<0.05, **= p <0.001

STUDY	N	R(psi_A, response_A)
Hypnosis	27	0.1730
Couples	54	-0.0377
Coffee	21	0.4277 *
Pooled	102	0.4051**

It should be noted that in all physiological data there are forward correlations between the physiological signals before and after the stimulus. For instance during an experiment the spontaneous activity may vary and periods with large variability will be interspersed with periods with lower variability across a trial. Thus a correlation analysis over trials will always produce a positive 'forward' correlation.

We tried to disentangle a potential retrocausal driven correlation from this normal correlation. What we did here is correlating the accuracy obtained using the pre-stimulus derived classifier with the accuracy obtained using the post-stimulus derived classifier. The situation then is less clear. One can argue that since the classifier is based upon EEG signals for which there are normal correlations, the classifiers themselves will be correlated and hence one should expect a correlation between the accuracies obtained using these classifiers. Currently we are breaking down these correlations for specific time intervals like (-600, -550) milliseconds, (-550, -500) milliseconds etc. We believe that such an analysis may provide unbiased answers to the question if we are dealing with (partial) time-symmetry in this kind of experiments.

Explorations

The current datasets allow for a large number of explorations such as:

- Where in the temporal or frequency domain is the predictive information?
- Where in the spatial (brain) domain 'originates' the predictive information?
- How is the predictive information distributed, are there many trials with a small amount of predictive information and a smaller number with a large amount?
- Is there time-symmetry?
- Is the predictive accuracy different for trials where the subjects did incorrectly 'miss' the face?

These explorations are currently underway and the results will be published elsewhere.

The Hypnosis Study

Only the hypnosis study, in spite of showing a result in the expected direction, turns out to be statistically non-significant. In at least one theoretical approach brain coherence is an essential factor and it is known that hypnosis results in dissociation of different parts of the brain and thus the effect of hypnosis might be that brain coherence decreases and hence the time-symmetry disappears.

There is a suggestion for this argument in the data. The very first block of trials in the hypnosis study is before any hypnosis takes place. That block shows a significant "retrocausal" effect with a predictive accuracy of 51.6% ($p<0.01$). The other blocks accuracies, after subjects are in the hypnotic state, are flat chance.

Alternative Normal Explanations for 'retrocausal' Results

Two of the three studies as well as the over-all analysis do show anomalous correlations between the brain signals and random future conditions. Before concluding that this is really anomalous, normal causal explanations have to be ruled out. We checked therefore possible patterns in the randomization of these future conditions, patterns that would allow the subjects to infer what the upcoming condition would be. No problems with randomization and also no signs of learning (and hence improving performance) were found. We also checked if the filter that was used to filter the raw EEG-data produced a response in the past as many digital filters do. Indeed we found that an impulse signal passing through the filter produced a signal that started about 40 milliseconds before the start of the original signal. However the interval that we used to 'predict' the upcoming stimulus condition was much earlier (-600 till -100 ms). To be sure, we also ran analyses using no filtering and no DC-detrend at all and results, although weaker, were still significant. There is of course a question if the assumed MCE of 50% accuracy is correct. As one of the reviewers put it: "…. Pattern classifiers may be over-generous…. (so have a MCE > 50%)". We therefore

repeated classification but we now used pseudo-labeled data, which should result in chance performance in absence of classifier bias. This procedure confirmed the MCE of 50%, as did the aforementioned control simulation with only noise data.

The fact that 2 of the 3 studies showed the same effect can be interpreted as a sign for the robustness of the effect. It could also be a sign that there is an undetected software problem. As an extra check we provided the raw data to an independent researcher who with his own software ran the same analysis and confirmed the obtained accuracies.

The robustness of the data is partly due to the number of trials in these studies. Generally, psi experiments do not use that many trials because there is a belief that boring experiments will not work. So 6-8 blocks of 120 trials per subject is really a lot. Just to illustrate: in Ganzfeld-telepathy experiments there is generally only one trial per subject.

But the effect size seems also to be rather large compared with other paradigms. We have wondered why this might be the case for these unselected subjects. A speculation is that this is due to the visual noise in which the face is embedded. This noise is dynamic (10 frames per second!). And there are signs that this dynamic visual noise induces a special alpha like brain. state. This speculation could be tested using different visual noise masks (dynamic and static).

Finally the apparent robustness of the results may be due to the fact that in this approach individual differences in EEG signals are not a source of uncontrolled variance. The analyses are done per subject. Idiosyncratic physiological behavior is taking care off by constructing specific templates per subject. An alternative way to construct the individual templates is to use neural network training. This has to be preceded by Independent Component Analysis because neural net training is inefficient if it is on the basis of highly dependent time-series. The advantage would be that interactions between channels in the predictive performance can be modeled.

PART II: THE GRONINGEN BILKING EXPERIMENT

Taken at face value we may conclude that in 2.7% (52.7-50%) of the cases the signal before the stimulus is 'determined' by the future condition.

What happens if we use this information to exclude the future that seems to be predicted? This is called bilking and it is one of the basic arguments used in the discussion on time-travel [17, 18].

Such an experiment apart from opening an avenue to testing potential time-loop paradoxes might shed some light on the theoretical models that have been proposed to account for the elusiveness of parapsychological results in the parapsychological literature.. In two of these models, the MPI model and the Generalized Quantum Theory model [19, 20] the claim is that no classical signals are allowed and any time one tries to use the anomalous correlation to transmit a classical signal, the correlations will disappear. In the CIRTS model] this restriction is weakened to the requirement that classical signals are not allowed only when they might be usable to create a paradox.

Global Modification to Paradigm to Test the Effect of Bilking

Because in this experiment we want to use the predictive information to manipulate the future, the bilking experiment consists of 2 stages.

Stage 1 (TRAINING): The training of the computer software to recognize the EEG patterns preceding specific future stimuli, in this case preceding the noisy face and preceding the noise-alone stimulus. This training stage becomes necessary because the bilking experiment hinges on the prediction of the future condition in real time and not in a post hoc way as we did in the 3 face experiments that we discussed in Part I.

Stage 2 (TESTING); The real time testing of the accuracy of this predictive pattern in terms of actual accuracies of the predictions in stage2 of the experiment. This accuracy is the dependent variable.

Each stage consists of a large number of trials. In stage 1 these are 250 trials and in stage 2 there are 750 trials. The task is extremely boring and according to parapsychologists totally unsuitable to assess psi effects.

Description of a Trial in the Bilking Experiment

A flow chart for the presentation of a trial is given in figure 3. Three types of trial can occur.

Normal trial (training-stage1 and test-stage2)

Like in the in Part 1 of this article reported face detection experiments, a normal trial consists of presentation of a 2 second visual stimulus while concurrently 8 channels of EEG are measured and processed. The visual stimulus consists of 21 (100 msec) frames of static noise. In the 11th frame a cartoon face may be superimposed on the noise. This is decided randomly (50% chance). The subject's task is to detect that face. (S)he indicates detection by a button press.

The Bilking trial (test-stage2)

In a normal trial the face condition is decided by a (quantum) random decision. I.e. this event is totally random. In the bilking condition the decision to show a face (or not) is totally algorithmic. What is happening in a bilking trial is the following: The software processes the EEG preceding the stimulus condition decision in real time and compares it with the predictive patterns assessed in stage 1. If the predictive pattern matches best with a future where a face will be presented the NO face will be shown and if the predictive pattern is closer to the pattern assessed in stage 1 for trials no face was show, then a face will be shown.

Three Conditions in Test-stage 2

In stage 2 there are trials in 3 conditions. Two of these conditions have been already described. This is the in the first place a normal trial (see above). Let call this the condition N. The second type of trial is the bilking trial (see above). Let's call this the condition B.

The third condition is introduced to have a handle to find out what the effect of the bilking condition might be. Actually this is a normal trial where the face-no face decision is not algorithmic as in the bilking condition but random. Let's call this condition N*. In fig. 3 it becomes clear why we label that condition slightly differently. The reason is that there is a different history of decisions leading to a N* trial than to a N trial.

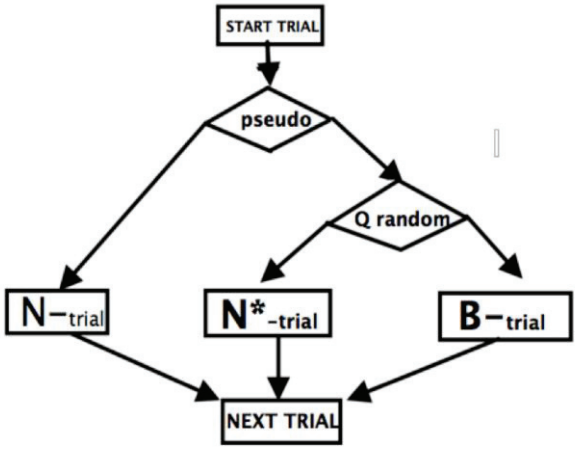

FIGURE 3.Decision tree to determine type of trial. The first decision is pseudorandom and split of a branch to a normal trial. The second decision is a quantum decision and splits into two branches, a normal trial and a bilking trial.

Very Preliminary Results

At the time of writing only 2 subjects have participated in this experiment. This is a far too low power to draw definitive conclusions, results are presented here for illustrative purposes.

Using the template that is constructed in the training stage 1, we find predictive accuracies below chance. For the normal (N) trials 42.67% and for the N* trials we find 40% accuracy. (For the B-trials the accuracy is of course 0 by design.) This strongly suggests that the predictive information in the templates constructed in the training stage was inferior to the information we generated by post hoc leave-one-out classification.

Therefore we also ran a classification procedure as described in part one, i.e. we are using the actual recorded data from stage 2 to construct the classification templates in a leave-one-out approach. This procedure allows us to draw conclusions about the actual information present in the signal. The prediction accuracies improve dramatically: for N it becomes 65.33%, for N* trials 65% and even in the bilking condition we get an above chance prediction accuracy of 56.36%.

New Dependent Variable

According to suggestion by York Dobyns (this volume) the QRNG deciding between N* and B trials should act in such a way to prohibit closed time like loops. This suggestion is based upon his reading of the analysis of closed time-like loops by Echeverria et al [17]. We assume that this implies that there will be bias such that there will be less bilking trials than can be expected. Preliminary results show the opposite, the QRNG is unexpectedly biased towards presenting bilking trials. Further control runs of the Random Number Generator are currently under way.

Apart from biasing the RNG there are other ways in which nature could prevent paradoxes. Especially the brain behavior of subject might be more noisy resulting in significant reduction of the accuracy of the predictive power. This could be visible in the EEG signals in the pre-stimulus period themselves or in the behavioral measures of the subject (how well he or she detects the face).

Planned Improvements

At the moment the predictive patterns for use in the bilking decision is determined in training-stage1. However if we apply the classification using a post hoc template as we did in the original analyses described in part 1, we find much better accuracies. This indicates that the predictive patterns assessed in the training are less efficient compared to the predictive patterns created post hoc using all the trials (except of course the trial where the future condition has to be predicted). This may be caused by two factors. In the first place in the three post hoc experiments we used much more trials to construct the predictive patterns. In the second place the predictive pattern thus constructed in the current bilking experiment is fixed and can't change over time. It could be that in test-stage2 the subjects gradually develop a different predictive EEG.

Thus an improvement would be to adjust the predictive patterns constructed in stage1 during stage2 by taking each new trial into account. This requires a substantial increase in computing power but it is not beyond the possibilities.

ACKNOWLEDGEMENTS

We thank Arnaud Delorme for independently checking part of the data and the analyses. We thank York Dobyns for pointing us to the Echeverrria et al article.

NOTE

Ethics restrictions prohibit public data sharing, but all data is available upon request from the authors, and will be deposited in the Heymans Institute Data Archive. The data-acquisition and the analysis software is available upon request, and will be published online via the Open Science Framework after publication.

REFERENCES

1. J. Mossbridge, P.E. Tressoldi and J. Utts, “Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis” Frontiers in Psychology, **3**, 390 (2012).
2. K. Popper, “The logic of scientific discovery”. (Routledge, 2005).
3. J. Hartwell, “Contingent negative variation as an index of precognitive information,” European Journal of Parapsychology, 83-103 (1978).
4. N. Don, B. McDonough and C. Warren, “Event related potentials (ERP) indicators of unconscious PSI: a replication using subjects unselected for PSI” Journal of Parapsychology, **62**, 127-145 (1998).
5. D.J. Bierman and J. Ditzhuijzen, “Anomalous slow cortical components in a slot-machine task”. Proceedings of the 49th Annual Parapsychological Association Convention. 5-19 (2006).
6. B. McDonough, N. Don and C. Warren, C. “Differential event related potentials in targets and decoys in a guessing task”. Journal of Parapsychology, **16**, 187-206 (2002).
7. T. Hintenberger, P. Studer, M. Jager, C. Haverty-Stacke and H. Walach, “The slide-show presentiment effect discovered in brain electrical activity”. The proceedings of the Parapsychological Association Convention, 1-14. (2006).
8. D. Radin and E. Lobach. "Toward understanding the placebo effect: investigating a possible retrocausal factor." The Journal of Alternative and Complementary Medicine **13**.7, 733-740 (2007):
9. D.J. Bierman, J.P. Spottiswoode, and A. Bijl. "Testing for Questionable Research Practices in a Meta-Analysis: An Example from Experimental Parapsychology." PloS one **11**.5 (2016): e0153049.
10. J. Jolij and M. Meurs. "Music alters visual perception." PLoS One 6.4 (2011): e18861. (2011).
11. J. Jolij, M. Meurs, and E. Haitel. "Why do we see what’s not there?." Communicative & integrative biology **4**.6: 764-767 (2011).
12. C. Bandt, M. Weymar, D. Samaga, and A.O. Hamm, A.O. A simple classification tool for single-trial analysis of ERP components. Psychophysiology, **46**, 747-757.(2009).
13. J. Jolij, H.S. Scholte, S. van Gaal, T.L. Hodgson and V.A. Lamme, “Act quickly, decide later: Long-latency visual processing underlies perceptual decisions but not reflexive behavior”. Journal of Cognitive Neuroscience, **23**(12), 3734-3745 (2012).
14. J. Jolij, “Social influence alters perception”. Perception, **41**, 209. (2012)
15. D.J. Bierman, "Anomalous switching of the bi-stable percept of a Necker cube: A preliminary study." Journal of Scientific Exploration, **25**.4 721 (2011)
16. D.J. Bierman, "Consciousness induced restoration of time symmetry (CIRTS): a psychophysical theoretical perspective." The Journal of Parapsychology, 74.2: 273 (2010).
17. F. Echeverria, G. Klinkhammer, and K.S. Thorne). “Billiard balls in wormhole spacetimes with closed timelike curves: Classical theory”. Physical Review D, **44**(4), 1077. (1991
18. Krasnikov, S. (2002). “Time travel paradox”. Physical Review D, **65**(6), 064013.
19. W. von Lucadou, “Some remarks on the problem of repeatability of experiments dealing with complex systems.” In: Information Dynamics (pp. 143-151). Springer US.
20. T. Filk and H. Römer, “Generalized Quantum Theory: Overview and Latest Developments”. (2012). http://arxiv.org/abs/1202.1659