Does Psychophysiological Predictive Anticipatory Activity Predict Real or Future Probable Events?

Patrizio E. Tressoldi, PhD# Massimiliano Martinelli, Dr, Luca Semenzato, Dr, and Alessandro Gonella, Dr

Background: The possibility of predicting random future events before any sensory clues by using human physiology as a dependent variable has been supported by the meta-analysis of Moss-bridge et al. (2012)[1] and recent findings by Tressoldi et al. (2011 and 2013)[2,3] and Mossbridge et al. (2014)[4] defined this phenomenon predictive anticipatory activity (PAA).

Aim of the study: From a theoretical point of view, one interesting question is whether PAA is related to the effective, real future presentation of these stimuli or whether it is related only to the probability of their presentation.

Methods: This hypothesis was tested with four experiments, two using heart rate and two using pupil dilation as dependent variables.

Results: In all four experiments, both a neutral stimulus and a potentially threatening stimulus were predicted 7–10% above chance, independently from whether the predicted threatening stimulus was presented or not.

Conclusion: These findings are discussed with reference to the "grandfather paradox," and some candidate explanations for this phenomena are presented.

Keywords: Random events, anticipation, prediction, pupil dilation, heart rate

(Explore 2015; 11:109-117 © 2015 Elsevier Inc. All rights reserved.)

INTRODUCTION

wThere is accumulating evidence that our nervous systems, the autonomic and the neurological, react to unpredictable (randomly presented) stimulation 3–10 seconds before they are triggered by a sensorial (visual or acoustic) stimulation. This anticipation is revealed by analyzing how psychophysiological signals change in relationship to the characteristics of future stimulations, for example, whether heart rate is enhanced before a future emotional stimulation compared to a non-emotional one. The anticipatory responses are analyzed averaging the psychophysiological responses, i.e., heart rate, skin conductance level, electroencephalography (EEG), etc., of all trials in order to extract the signal from the noise. A prototypical response is presented in Figure 1.

The possibility of predicting random future events using human physiology as dependent variable, before any sensory clues, is now supported by the meta-analysis of Mossbridge et al.,[1] reporting an estimated effect size of 0.21, 95% confidence interval (CI) = 0.13–0.29. This phenomenon, was defined predictive anticipatory activity (PAA), and its possible mechanisms, the theoretical implications, and its potential practical applications are discussed by Mossbridge et al.[4]

Dipartimento di Psicologia Generale, Università di Padova, via Venezia 8, Padova 35131, Italy

 # Corresponding author.
e-mail: patrizio.tressoldi@unipd.it

The interpretation of this apparent violation of time-symmetry is still under theoretical and empirical investigation. Given the aim of this article, we will discuss them only briefly in the discussion.

Tressoldi et al.[2,3] started a line of research aimed at using PAA to predict the category (neutral vs. emotional) of each stimulus presented randomly, at the level of single trials. In these experiments, there were two main methodological differences with respect to the typical procedure used to study the PAA. The first is that PAA is not averaged among all the trials of the experiment but used at the level of single trial to predict the category of the future events. It is clear that this implies a higher difficulty extracting the signal from the noise. The second main difference is that the aim of this line of research was not only to see whether the PAA mirrored the physiological reactions observed after the stimuli presentation but also to see whether the accuracy of these predictions was above the expected chance, for example, 50% when there are two categories to predict. A strong demonstration that PAA can predict future random events well above the chance will open the door to implement practical applications.

In studies by Tressoldi et al.,[2,3] it was shown that pupil dilation (PD) PAA predicted the random presentation of a neutral (a neutral sound or a smile) or an alerting stimulus (an alerting sound or an image of a gun associated with an acoustic shot) 6–10% above the chance expectation of 50%.

One interesting question is whether this PAA is related to the effective, real future presentation of these stimuli or whether it is related only to the probability of their presentation.

ISSN 1550-8307/$36.00

http://dx.doi.org/10.1016/j.explore.2014.12.003

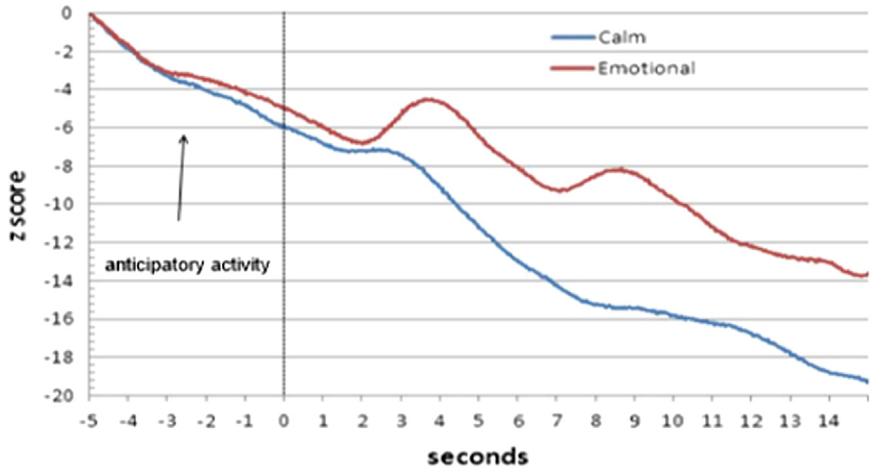

Figure 1. Example of a predictive anticipatory response related to two future events of different emotional content.

From a philosophical point of view, this problem is known as the "bilking argument" first introduced by Max Black.[5] The "bilking argument" states that if B is earlier than A, and let B be the alleged effect of A, if we assume that A causes B even though A is later than B, it is possible, in principle, to intervene in the course of events and prohibit A from occurring. But if this is the case, A cannot be the cause of B; hence, we cannot have backward causation or anticipatory prediction. Another name for this problem is "grandfather paradox"[6] described as follows: "*The time traveller goes back in time and kills his grandfather before his grandfather meets his grandmother. As a result, the time traveller is never born. But, if he was never born, then he is unable to travel through time and kill his grandfather, which means the traveller would then be born after all, and so on*."

One solution to this conundrum is to devise experiments where the predicted stimulus is not presented but skipped or deleted. If the predicted event is skipped, it cannot exert any backward effect and hence the prediction accuracy should be at chance. If, however, the prediction accuracy is above chance, it is necessary to explain which sort of information can be used to predict an event that never happened. We postpone the discussion of this problem to the end of the presentation of the results of all four experiments.

EXPERIMENTS 1 AND 2

The first two experiments are conceptual replications of studies by Tressoldi et al.,[2,3] using heart rate (HR) as PAA, instead of PD. In the first experiment, all future random events will be presented. In the second one, predicted alarming events will be skipped.

Method

Participants. Estimating an effect size of approximately 0.30 as observed in Tressoldi et al.,[2,3] to achieve a statistical power above 0.80, setting $\alpha = 0.05$, an opportunity sample of 100 students and personnel from Padova University[7] were recruited by a research assistant to participate in an experiment on a gambling task. The final sample comprised 28 males and 72 females, with age ranging from 23 to 35 years. Their participation was compensated with €5.

Ethics statement. Participation inclusion followed the ethics guidelines in accordance with the Helsinki Declaration, and the study was approved by the Ethics Committee of Dipartimento di Psicologia Generale, the hosting institution. Before taking part in the experiment, each participant provided written consent after reading a brief description of the experiment.

Apparatus and procedure. HR was detected by a photoplethysmograph connected to the index finger of the left hand.[8] The signal was subsequently conveyed to a Pulse Monitor 701 and to a Metex 3850 D digital multimeter and fed to a PC for online data acquisition. The software for HR data acquisition, visualization, and its connection with the presentation of the two sounds was developed by the two co-authors, M.M. and L.S., in E-Prime™ v.2.0. Heart rate per minute was automatically estimated using the formula *P* (pulse) = 60,000 ms/inter beat interval (IBI).

The two sounds (available in Ref. 9) were chosen from the International Affective Digitised Sounds (IADS) collection,[10,11] selected to trigger an alert or a neutral reaction.

The procedure comprised two phases, a preliminary and an experimental one. The preliminary phase was devised only to familiarize the participants with the procedure. Participants were required to sit in a comfortable chair in a light- and sound-attenuated lab, facing a PC monitor. After they had been connected with the photoplethysmograph, they were instructed not to move their body, breathe regularly avoiding deep breaths, wear the headphones (model Inno Hit SH-154), and control on the PC monitor if their HR proceeded regularly. When the research assistant was certain that the HR was stabilized, that is, the HR values varied smoothly without peaks, he started the software that controlled the random presentation of the two sounds after the participant pressed the central key of a response-box (Figure S3). The session ended after the presentation of the two sounds 10 times each. The choice to use a fixed number of data instead of the data recorded within a fixed time window allowed the individual differences to be taken into account and the average HR to be calculated using the same number of data.

Experimental phase. The experimental phase comprised two sessions, one we defined "intentional," where the participants were required to predict the key associated with the neutral sound and avoid the alarming one by using their intuition. In the second one, their HR PAA was used to estimate the prediction of the future events.

Each participant was instructed to monitor their HR on the monitor. When the HR looked stabilized, the participant had to press the central key of the response-box, colored yellow, triggering the software, which collected 10 anticipatory HR samples. Immediately after, he/she had to choose the left or right response-box colored red and associated with the neutral sound. When the software showed the go signal on the monitor, he/she had to press the chosen response-key. The neutral sound was associated with the left or right response-box using the E-Prime™ random routine after each trial reproducing a sequence with replacements. The session ended after 20 trials.

Prediction of the HR PAA. The prediction of the HR PAA was calculated offline applying the following algorithm: in order to take into account the individual differences, we standardized the HR values related to the 20 trials measured in the anticipation phase to *z* scores for each participant, and the corresponding means associated with the two stimuli were calculated. In this way, one of the mean was always above zero and the second one below zero, except when they were identical up to the third decimal, a condition that never happened. The prediction for each trial was obtained simply by defining whether the value of HR, above or below zero, corresponded to the future sound that was chosen randomly. For example, if the HR standardized means associated with the neutral and alarming sounds were 0.25 and −0.15, respectively, each HR value above zero predicted a neutral sound and each value below zero predicted an alerting sound. At the end of the trial, the sum and the percentage of hits (correct predictions) were calculated for each participant. Overall, 56% of participants showed a higher HR associated with the alarming sound.

Statistical methods. In all experiments, we will use both a frequentist parameters estimation and a Bayesian model comparison approach, according to the American Psychology Association,[12] Kruschke,[13] and Wagenmakers et al's.[14] statistical recommendations.

This statistical approach is recommended to limit the shortcomings of the classical Null Hypothesis Significant Testing.[15] Basically, each parameter of interest (mean, correlation, etc. and effect size) will be estimated for its precision by the confidence intervals. For those interested in the classical statistical significance with this approach, it is sufficient to check if the confidence intervals include (not significant) or exclude (significant) zero.

Inferential frequentist estimates will be applied both to the sum and the means of correct guesses (hits) using a binomial and a one-sample *t*-test statistical test, respectively. Confidence intervals will be estimated using a bootstrap procedure based on 5000 samples.

Bayesian statistics. We will adopt a model comparison approach contrasting the alternative hypothesis of a higher difference with respect to the mean chance expected (MCE) (H1) with the Null Hypothesis (H0) of a zero difference with respect to the MCE. We will calculate the Bayes Factor ($BF_{H1/H0}$) using the software implemented by Morey and Rouder[16] for the comparison with the one-tailed one-sample *t*-test, applying Jeffreys, Zellner, Siow (JZS) prior[17] setting an effect size of 0.3, as suggested by Rouder et al.[18] The JZS prior represents a combination of the Cauchy distribution and the probability of variance equals to $1/\sigma^2$.

Expectation bias control. Expectation bias is related to the human propensity to expect a "tail" in a coin toss after observing a series of "head" outcomes (the gambler's fallacy). The reason expectation bias can potentially explain PAA is that a series of (randomly selected) neutral stimuli may produce a physiological shift toward excitement as the presumably imminent emotional trial approaches. In a sequence of trials with several such series of neutral events preceding emotional events, simulations suggest that the resulting physiological data could mimic a PAA effect.[19,20] Thus, to understand the mechanisms underlying PAA, it is crucial to determine for each experiment whether expectation bias was a potential explanation for the reported outcome.

Results

The mean of the correct predictions of the two sounds in the "intentional" session was below the MCE: 0.43 for the alarming sound (95% CI: 0.37, 0.48) and 0.44 for the calm sound (95% CI: 0.38,0.49). In Table 1, we report the descriptive and inferential statistics related to the prediction accuracy using the HR PAA.

Expectation bias. The results are presented in Figure S1 where the mean accuracy percentages related to the number of repetitions of the same type of stimulus are shown. For example, in repetition 1, the smile was presented once and after the alarm followed, whereas in repetition three, the smile was presented three times consecutively and the alarm followed. If the participants had adopted the "gambler fallacy" heuristic, we should have expected an almost linear decrease of their accuracy with the number of repetitions of the same type of stimulus. On the contrary, if they adopted

Table 1. Estimates of the Percentages of Correct Prediction of the Two Sounds and Their Estimation With Respect to the Mean Chance Expected, 0.50 and Corresponding Bayes Factors of experiment 1

Sound	Mean and 95% CIs	SD	ES and 95% CIs	$BF_{(H1/H0)}$[a]
Alarm	0.56 (0.53, 0.58)	0.15	0.38 (0.18, 0.58)	4.5
Neutral	0.58 (0.55, 0.60)	0.12	0.68 (0.46, 0.90)	15.9

ES: effect sizes, SD: standard deviation.
[a]Setting effect size, $r = 0.3$.

Table 2. Estimates of the Percentages of Correct Prediction of the Two Sounds and Their Estimation With Respect to the Mean Chance Expected, 0.50 and Corresponding Bayes Factors of experiment 2

Sounds	Mean and 95% CIs	SD	ES and 95% CIs	$BF_{(H1/H0)}$ [a]
Alarm	0.56 (0.53, 0.58)	0.13	0.45 (0.25, 0.65)	6.7
Neutral	0.57 (0.54, 0.59)	0.13	0.52 (0.31, 0.73)	9.3

[a]Setting effect size, $r = 0.3$.

the "hot hand" heuristic, we should expect an almost linear increase of their accuracy in predicting the smile with the increase of its repetitions.

The slope values are 0.004 (95% CI: −0.025, 0.03) and −0.027 (95% CI: −0.057, 0.007) for the alerting and the neutral sounds, respectively, showing a non-statistically significant small expectation bias, favoring the prediction of the alarm sound and decreasing the prediction of the neutral sound.

EXPERIMENT 2

As anticipated, this experiment is identical to Experiment 1, except the skip of the alerting sounds when predicted (see the section Procedure).

Method

Participants. The same number of participants of the previous experiment were recruited using the same procedure.

Apparatus and procedure. The apparatus was the same as the previous experiment. The procedure was changed as follows. In the preliminary phase, the mean HR associated with the presentation of two sounds was calculated for each participant serving as an individual baseline for the experimental phase.

Participants were instructed to perform the task as in Experiment 1. However, before the go signal, the software compared the mean of the HR PAA with the two means related to both the alerting and the neutral sound recorded in the preliminary phase. If the difference between the average HR PAA and the HR average related to the alarm sound recorded in the preliminary phase was lower than the difference related to the neutral sound, the software skipped the presentation of the alarm sound. For example, if the overall average HR PAA was 80, and the differences with the average HR related to the presentation of the alarm and the neutral sound recorded in the preliminary phase were five and eight, respectively, the software predicted the presentation of an alarm sound skipping its presentation. In other words, if the HR PAA was closer to the HR related to the alerting sounds measured in the preliminary phase, the software predicted the presentation of this type of sound, blocking its presentation. On the contrary, if the HR PAA was closer to the HR related to the neutral sounds measured in the preliminary phase, the software predicted the presentation of this type of sound that was delivered. In summary, the alarm sound was presented only when the HR PAA erroneously predicted a neutral sound.

Results

In the "intentional" sessions, the means of predictions related to the two sounds were within the MCE: 0.43 for the alerting sound (95% CI: 0.45, 0.53) and 0.47 for the neutral sound (95% CI: 0.42, 0.50). In Table 2, we report the descriptive and inferential statistics related to the prediction accuracy using the anticipatory HR activity.

Expectation bias. The results are presented in Figure S2. The slope values are 0.021 (95% CI: −0.007, 0.05) and 0.001 (95% CI: −0.03, 0.03) for the alerting and the neutral sounds, respectively, showing a small non-statistically significant expectation bias, favoring the prediction of both sounds.

Comparison with experiment 1. A direct comparison between the results of the two experiments is presented in Figure 2 where it results an almost identical percentage of correct predictions of the two sounds independently from the feedback condition (both sounds presented in the Experiment 1, skipping of the alerting sounds in Experiment 2 when predicted by the software).

EXPERIMENTS 3 AND 4

The following two experiments are a variant of the experiments of Tressoldi et al.[3] The only difference being that the negative event predicted in the anticipatory phase was skipped instead of presented. Comparing the results with the previous experiments and the following ones, it is possible to test further the "bilking paradox," that is, whether it is possible to avoid predicted future negative events, giving more support to the results observed in the experiment 2.

Method

Participants. As in Tressoldi et al.[2] Experiment 1, an opportunity sample of 80 students and personnel from Padova University were recruited by a research assistant to participate in an experiment on a gambling task. The final sample comprised 40 males and 40 females with a mean age of 29.3 and with a standard deviation of 3.8. Participants' participation was compensated with €5.

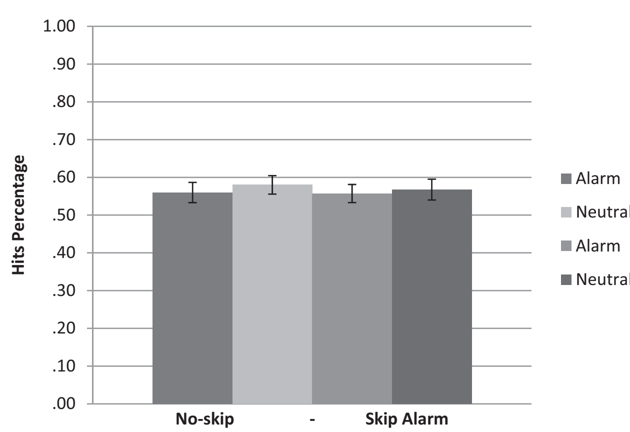

Figure 2. Averaged accuracy percentages among all participants related to the prediction of the two sounds, with corresponding CIs, observed in Experiment 1, No-skip and Experiment 2, Skip Alarm.

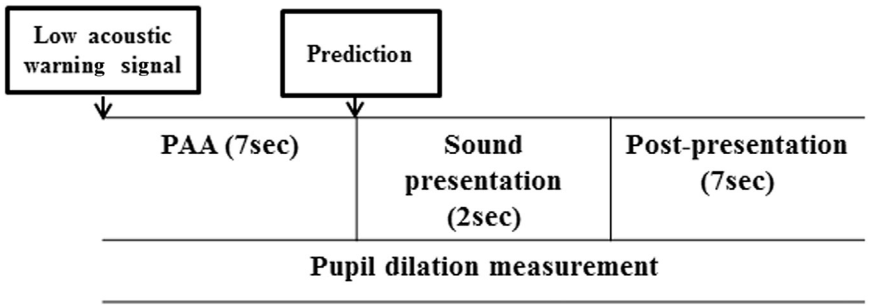

Figure 3. Sequence of events in the experimental session.

Procedure. All participants were tested in a sound- and light-attenuated laboratory located in the Dipartimento of Psicologia Generale of Padova University.

The procedure comprised two sessions.

Baseline session. During this session, the participants were requested to listen passively to two sounds (an alerting and a neutral sound) and to look inside a white circle presented in the middle of the monitor to allow the Tobii™ 120 eye-tracker to record their PD.

Sounds were conveyed to participants by headphones (model Inno Hit SH-154), following a random sequence and inter-stimulus intervals ranging from one to three seconds. After this phase, which lasted for no more than three minutes, the average PD related to the neutral and the alerting sound was calculated and stored to be used in the experimental session.

Experimental session. In this session, a special devised prediction algorithm was used to predict the upcoming sound and skip the alarming one.

Prediction algorithm. The prediction algorithm is quite simple (see sequence of events in Figure 3). A software subtracted the mean PD PAA of each participant measured in the anticipatory period, from each of the two PD means related to alerting and neutral sounds recorded in the baseline session. The comparison with less difference was used to predict the category of the sound to be delivered. For example, if the average pupil dilation for alerting and neutral sounds for participant *X* measured in the baseline session was 3.5 mm and 4 mm, respectively, and the PD measured in the anticipatory period was 3.4 mm, the algorithm would predict an alerting sound.

Different from Tressoldi et al.[1] Experiment 1, in this case, if the algorithm predicted an alerting sound, the software automatically skipped its presentation and the participants did not hear anything. With a perfect, 100% prediction accuracy, participants could hear only the neutral sounds.

Table 3. Descriptive Statistics Related to the Count and the Mean of Hits of experiment 3

	Neutral Sounds	Alerting Sounds
Hits %	0.43 (347/800)	0.58 (463/800)
M	43.4	57.9
SD	30.1	30.4

Preliminary data analysis. Before proceeding with the statistical analyses, the data for each participant were screened for artifacts. All artifacts, i.e., missing or anomalous (PD values close to/below 1, or above 10) data recordings related to PD easily detected by inspecting the raw scores saved in the individual files, were eliminated. If they exceeded the threshold of 60%, that is, 12 out 20 trials, the entire participant was excluded and substituted to keep the total sample equal to 100. The overall percentage of artifacts was 4%. Blinks are automatically detected by the eye-tracker and deleted from the statistical analyses.

Results

In Table 3, we report the descriptive statistics and in Table 4, the inferential ones based on the frequentist parameter estimation and the Bayesian model comparison.

Expectation bias control. The average accuracy in the prediction of the neutral and alerting sounds at different lags (sequences of repeated sounds) is represented in the Figure S4.

The slope values are −0.008 (95% CI: −0.04, 0.03) and −0.03 (95% CI: −0.06, 0.02) for the alerting and the neutral sounds, respectively, showing a non-statistically significant small decrement in the prediction accuracy for both sounds. It seems then that the expectation bias, if present, harms the prediction accuracy.

Comparison with the experiment 1 in Tressoldi et al.[2] In Figure 4, we show the average hit percentages obtained by all participants related to both neutral and alerting stimuli in the skip and no-skip conditions.

Table 4. Inferential Statistics of experiment 3

	Neutral Sounds	Alerting Sounds
One-sample *t*-test	−1.96	2.31
ES (95% CI)	0.22 (0.0, 0.44)	0.26 (0.04, 0.48)
$BF_{H1/H0}$	0.22	0.85
Binomial *z* test	−3.7	4.42
ES (95% CI)	0.41(0.18, 0.64)	0.49 (0.26, 0.72)
$BF_{H1/H0}$	4.2	6.5

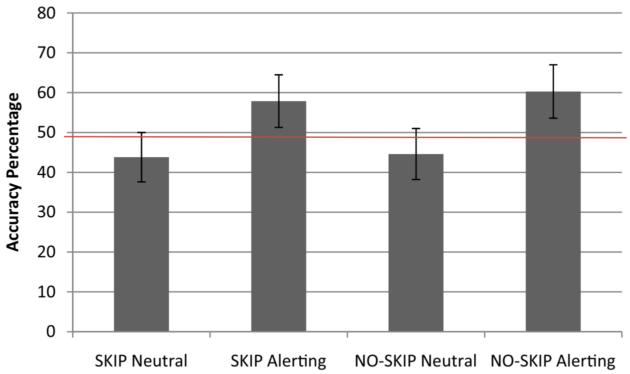

Figure 4. Averaged accuracy percentages among all participants related to both neutral and alerting stimuli in the skip and no-skip conditions. Error bars are 95% confidence intervals.

Comment. The results clearly show no difference between the conditions when predicted sounds are presented or when they are skipped.

Study limitations. As indicated in Tressoldi et al.,[2] the choice to present the sequence of sounds without replacement introduces a bias because there is a small probability (2.71e-5) that participants can predict the sound category using a strategy of counting the number of sounds of each category above the level of chance. When all sounds of one category are presented, the remaining ones are clearly exemplars of the second category, a strategy requiring a very high cognitive load.

EXPERIMENT 4

This is an exact replication of the experiments by Tressoldi et al.,[3] aimed at testing if the observed prediction accuracy holds even when the alerting stimuli get skipped when predicted from the measurement of the PD before their presentation.

Method

Participants. Despite the relatively large effect size of 0.40 observed in the original study, we decided to recruit again 100 participants to keep the number of participants identical.

Materials and apparatus. The two target stimuli, the gun and the smile, and the door, were calibrated for luminance (300 × 471 pixel; 72 horizontal and vertical dpi). The door was colored in black similar to the video background to avoid PD modification consequent to differences in luminosity. Their luminance measured using cd/m^2 with a Minolta® photometer was, for the gun: 15 center, 90 periphery, Smile: 73 center, four periphery, and door: 48 center, eight periphery, taken at 50 centimeters from the monitor.

Eye-Tracker Apparatus: The eye-tracker model Tobii T120® has the following technical characteristics: data rate, 120 Hz; accuracy, 0.5°; freedom of head movements, 30 × 22 × 30 cm; monitor, 17 in; 1280 × 1024 pixels; and automatic optimization of bright-dark pupil tracking. PD is measured automatically in millimeters by the apparatus using the incorporated near-infrared detectors and software. Blinks were automatically detected and missed for the PD averaging.

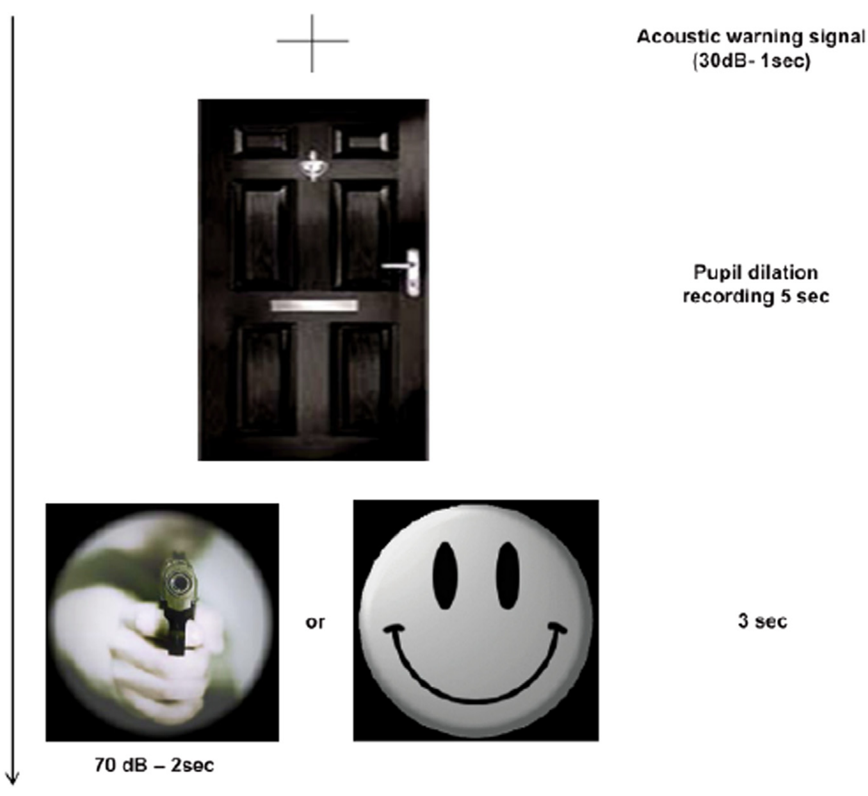

Figure 5. Sequence of events of the experimental phase of experiment 4.

Table 5. Descriptive Statistics of Hits Percentage for the Two Stimuli and Overall of experiment 4

		Smile	Gun	Overall
Hits	Mean %	0.577	0.587	0.582
	SD	0.15	0.16	0.09
	Sum	533/943	550/960	1083/1904

These data were fed to an original software for their storage. This program, created using E-Prime™ v.2.0, written by two of the authors (M.M. and L.S.) and interfaced with the eye-tracker, controlled events presentation and pupil size automatic recording.

The sampling with replacement of the two stimuli in the two series of 10 trials was randomized using E-Prime™ v2.0 randomized statement and Random function, which was reset after every trial.

This procedure guarantees against the possibility to guess the incoming stimulus by learning implicit and explicit rules.

The light in the laboratory was constantly dim, approximately 30 cd/m^2, to avoid undesired or unrelated changes to the participants' pupils.

Procedure. We used the same procedure used by Tressoldi et al.[3] with only the following modification: when the PD PAA was more similar to the PD related to the gun stimulus observed in the preliminary phase, the software skipped the presentation of the gun and the associated sound. For example, if in the preliminary phase the PD was on average 4.0 and 3.5 mm. for the gun and the smile, respectively, and in the anticipatory phase the average PD was 4.3, their difference were 0.3 and 0.8. In this case the gun presentation was skipped. With this procedure, the gun presentation was skipped in 86.3% of cases.

Preliminary phase. This phase served to familiarize the participants with the procedure and record the PD related to the presentation of the two images. The sequence of events is shown in Figure 5. A moderate acoustic signal was delivered simultaneously to a cross in the middle of the screen for one second to warn the participant that the trial was to start. Just after, a black door was presented in the middle of the screen for five seconds followed by a picture of a smile or of a shooting gun associated with a sound reproducing a gunshot presented in random order while their PD was recorded. After an intertrial period ranging randomly from 5 to 10 seconds, a new trial was delivered for a total of 20 trials. Participants were instructed to simply watch what would happen on the screen of the pc monitor.

Experimental phase. When the black door was presented in the middle of the screen, the PD PAA was recorded for the analysis of the prediction accuracy.

Individual prediction accuracy. We used the same prediction algorithm used in Experiment 2: in order to take into account individual differences, we standardized the PD values related to the 20 trials measured in the anticipation phase to z scores for each participant, and the corresponding means associated with the two stimuli were calculated. In this way, one mean was always above zero and the second one below zero except when the means of the two stimuli were identical up to the third decimal, a condition that never happened.

The prediction for each trial was obtained simply by defining whether the value of PD, above or below zero, corresponded to the stimulus that was chosen randomly. For example, if the PD standardized means associated with the smile and the gun were 0.25 and −0.15, respectively, each PD value above zero predicted a smile and each value below zero predicted a gun. At the end of the trial, the sum and the percentage of hits (correct predictions) were calculated for each participant. Overall, 54% of participants showed a higher PD PAA for the smile image.

Results

Overall prediction accuracy. In Table 5, we report the descriptive statistics; in Figure 6, the average hit percentages obtained by all participants with their 95% CIs associated with the results of the replication experiment reported by Tressoldi et al.[3]; and in Table 6, the effect sizes estimation and the $BF_{H1/H0}$ of the two stimuli and overall with respect to the MCE and the corresponding 95% CIs.

Expectation bias control. The average accuracy in the prediction of the neutral and alerting sounds at different lags (sequences of repeated stimuli) is represented in the Figure S4.

The slope values are 0.011 (95% CI: −0.017, 0.04) and −0.013 (95% CI: −0.04, 0.015) for the smile and the gun, respectively, showing a non-statistically significant small increment in the prediction accuracy of the smile and the contrary for the gun.

DISCUSSION

The main question addressed with these series of experiments was whether the PAA accuracy varies depending on whether the potentially alarming or threatening future event is presented or skipped.

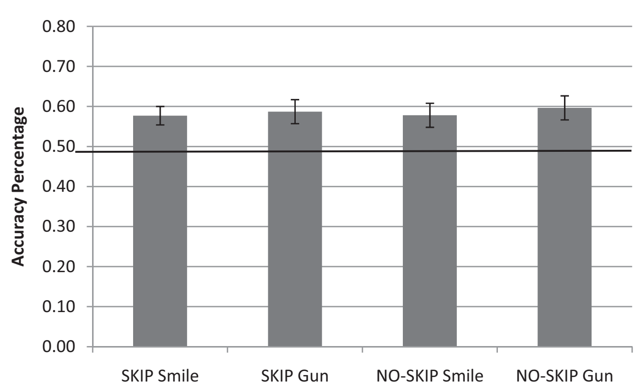

Figure 6. Averaged accuracy percentages among all participants related to the prediction of the two images, with corresponding CIs.

Table 6. Inferential Statistics: Effect Sizes With 95% CIs and $BF_{H1/H0}$ Values of Hits Percentage for the Two Stimuli With Respect to the Mean Chance Expected, 50% of experiment 4

	Smile	Gun	Overall
Binomial *z*	3.97	4.49	5.98
$ES(z/\sqrt{n})$	0.39 (0.19, 0.59)	0.49 (0.28, 0.70)	0.59 (0.38, 0.80)
One-Sample *t*	5.3	5.5	9.3
$ES\ d\ (t/\sqrt{n-1})$	0.53 (0.32, 0.74)	0.55 (0.34, 0.76)	0.93 (0.69, 1.16)
$BF_{H1/H0}$	9.8	10.5	27.7

The results of the four experiments give converging support to the hypothesis that PD and HR PAA can predict not only real future random events but probable ones too. The hit percentages, ranging approximately from 6% to 10% above the MCE, are almost the same between the conditions when the predicted stimulus is presented and those when one of them, the potentially threatening one, is skipped.

The results observed in experiments 2, 3, and 4 exclude the possibility that the participants could have adopted a non-conscious strategy to avoid upcoming alerting stimuli altering their HR or PD to match the alerting stimuli. If they adopted this strategy, we should have observed accuracy percentages close to 100% for this type of stimuli and accuracy percentages close of below chance for the non-alerting stimuli.

If this phenomenon is confirmed by independent replications, it will raise more questions than answers with respect to the processes regulating it. If the PAA was only a sort of anticipation of future real events, even if random, we could postulate that our psychophysiological system can anticipate random future events by a sort of biological entanglement in time mechanism similar to that observed by Atmanspacher and Filk[21] in the perception of ambiguous images. However, if it is possible to avoid the perception of anticipated events, it could give rise to a sort of "grandfather paradox," thus incurring the "bilking argument."

One possibility to solve these contradictions is to postulate that they are valid only if time is regulated strictly by deterministic, causal laws, but not if it is regulated by probabilistic laws. In this case, the possibility to change the future is allowed because it is not fixed. If this is so, this situation does not allow for the possibility of obtaining perfect predictions. Our results seem to be in agreement with this interpretation.

Acknowledgments

We acknowledge the accurate review and comments of one anonymous reviewer, which allow us to improve the previous version of the article. English was revised by the professional service Proof Reading Service.

APPENDIX A SUPPLEMENTARY MATERIAL

Supplementary data are available in the online version of this article at http://dx.doi.org/10.1016/j.explore.2014.12.003.

REFERENCES

1. Mossbridge J, Tressoldi P, Utts J. Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. *Front Psychol.* 2012;3:390. http://dx.doi.org/10.3389/fpsyg.2012.00390.
2. Tressoldi PE, Martinelli M, Semenzato L, Cappato S. Let Your Eyes Predict prediction accuracy of pupillary responses to random alerting and neutral sounds. *SAGE Open*. 2011;1(2):1–7.
3. Tressoldi PE, Martinelli M, Semenzato L. Pupil dilation prediction of random events [v2; ref status: approved with reservations 2, http://f1000r.es/3dw.]. *F1000 Res.* 2014;2:262. http://dx.doi.org/10.12688/f1000research.2-262.v2.
4. Mossbridge JA, Tressoldi P, Utts J, Ives JA, Radin D, Jonas WB. Predicting the unpredictable: critical analysis and practical implications of predictive anticipatory activity. *Front Hum Neurosci*. 2014;8:146. http://dx.doi.org/10.3389/fnhum.2014.00146.
5. Black M. Why cannot an effect precede its cause. *Analysis*. 1956;16:49–58.
6. Deutsch D, Lockwood M. *The quantum physics of time travel. Science Fiction and Philosophy–From Time Travel to Superintelligence.* NY, USA: John Wiley & Sons; 2009;322–334.
7. Faul F, Erdfelder E, Lang AG, et al. G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175–191.
8. Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas*. 2007;28(3):R1–R39.
9. Tressoldi P. *PredictiveAnticipatoryActivity*. *figshare*. ⟨http://dx.doi.org/10.6084/m9.figshare.978510.⟩; 2014.
10. Bradley MM, Lang PJ. *International affective digitized sounds (IADS): stimuli.Instruction Manual and Affective Ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida; 1999; (Technical Report B-2).
11. Bradley MM, Lang PJ. Affective reactions to acoustic stimuli. *Psychophysiology*. 2000;37(2):204–215.
12. APA Publications and Communications Board Working Group on Journal Article Reporting Standards. Reporting standards for research in psychology: why do we need them? What might they be? *Am Psychol*. 2008;63(9):839–851.
13. Kruschke J. Introduction to special section on Bayesian data analysis. *Perspect Psychol Sci*. 2011;6(3):272–273.
14. Wagenmakers EJ, Wetzels R, Borsboom D, et al. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem 2011. *J Pers Soc Psychol*. 2011;100(3):426–432.
15. Tressoldi PE, Giofré D, Sella F, Cumming G. High impact = high statistical standards? not necessarily so. *PLoS ONE*. 2013;8(2):e56180.
16. Morey RD, Rouder JN. *Package Bayes factor*. ⟨http://bayesfactorpcl.r-forge.r-project.org⟩; 2013.
17. Jeffreys H. *Theory of Probability*. Oxford, England: Oxford University Press; 1961.

18. Rouder JN, Speckman PL, Sun D, et al. Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev*. 2009;16(2):225–237.
19. Dalkvist J., Westerlund J., Bierman D. *A computational expectation bias as revealed by simulations of presentiment experiments*. Proceedings of the 45th Annual Convention of the Parapsychological Association. Paris; 2002: 62–79.
20. Wackermann J. *On cumulative effects and averaging artefacts in randomised S-R experimental designs*. Paper Presented at the 45th Annual Convention of the Parapsychological Association. Paris; 2002: 293–305.
21. Atmanspacher H, Filk T. A proposed test of temporal nonlocality in bistable perception. *J Math Psychol*. 2010;54(3): 314–321.