Skin Conductance Prestimulus Response: Analyses, Artifacts and a Pilot Study

by

S. James P. Spottiswoode, B.Sc & E. C. May, Ph.D.
Laboratories for Fundamental Research
Palo Alto, CA

Abstract

Previous studies have suggested that the human autonomic nervous system responds to stimuli 2-3 seconds before presentation. In these studies randomly chosen photographs with high and low affectivity were presented to participants. Ensemble averaging of skin conductance in the prestimulus epochs showed a differential response between high and low affectivity photographs. In our protocol the problem of idiosyncratic responses to pictorial stimuli was avoided by using audio startle stimuli. Stimulus type was determined just before presentation by a true random generator. Participants heard 20 stimuli per session with a 50% chance of an audio startle as against a silent control. Our dependent variable was the proportions of 3-second epochs prior to audio and control stimuli in which a skin conductance response, that is a minimum in skin conductance followed by a maximum, occurred. We found a significant effect ($N = 125$, Z-score (Z) = 3.27, effect size (ES) = 0.0901 ± 0.0275, $p = 5.4\times10^{-4}$). Explanations for this result as an artifact were examined and rejected. We show that a significant result from an average-based epoch analysis in this type of experiment is not a necessary requirement to demonstrate significant evidence for a prestimulus response.

Background

While the autonomic nervous system response to emotional stimuli is long established (Andreassi, 1989; Bouscein, 1992), recent research by Radin (1997a,b), Bierman & Radin (1997, 1998), and Bierman (2000) has suggested that skin conductance and other autonomic measures can act as a statistical predictor of a future experience. These results have begun to attract interest elsewhere (Parkhomtchouk, et al, 2002) and are beginning to be investigated with other techniques such as functional magnetic resonance imaging (Bierman & Scholte, 2002).

To illustrate their method, imagine a large pool of emotional photographic stimuli that have been previously evaluated with regard to the degree of emotional content. Schematically, the protocol in Radin's (1997b) experiments was as follows:

- A participant pressed a button to initiate a trial while watching a blank computer display.

- After a predefined time called the prestimulus period, a selection of a single photograph was made at random and was displayed for three seconds. The ratio of calm to emotional targets was 2:1, as the target pool had twice as many calm as emotional photographs.
- The trial ended after a 15 second post stimulus period.

The above sequence was repeated 30 times. On the average, there were 10 emotional photographs and 20 calm photographs. Using the change of skin conductance from the trial onset as the dependent variable, ensemble averages were constructed separately for the emotional and calm presentations. The difference between the ensemble averages in the prestimulus region was the session dependent variable. Relevant statistics were determined by randomized permutation techniques.

We offer criticism of this approach on two grounds:

1. Photographic stimuli can elicit idiosyncratic responses. For instance, a picture that has been rated as having a low average affectivity may have, for some individual participants, a large affectivity. This mechanism reduces the contrast between arousing and calming presentations and constitutes an unwanted source of variance in these designs.
2. Average-based epoch analyses, which where used in these studies, are sensitive to amplitude outliers, even with a randomized permutation analysis test statistic.

Critchley et al. (2000) have shown using fMRI techniques that there are three components that contribute to an overall skin conductance response (SCR). These are responses to an emotion, an external stimulus and a motor movement. If photographic stimuli, such as those contained in the Lang et al. (1999) International Affective Picture System (IAPS) set, are used in a prestimulus response study the first two of the three sources of skin conductance changes are involved and the emotional component can introduce variance due to idiosyncratic responses. Thus, for the current experiment we have abandoned the IAPS stimuli in favor of an audio startle stimulus comprised of a 1 s burst of 97 db white noise delivered to sound-isolating earphones.

Hypothesis

The level of arousal of the autonomic nervous system, as measured by changes in skin conductance, responds in advance significantly more before future audio stimuli than before future silent control stimuli.

Method

We used an SC5-SA skin conductance monitor, manufactured by Contact Precision Instruments. This unit is specified to have an absolute accuracy of ± 0.1 μSiemens, a DC excitation voltage of 0.5 V, a constant sample rate of 40 samples s^{-1}, and a relative

accuracy of 5.96×10^{-6} μSiemens. The unit contains a hardware low-pass filter with an upper cutoff frequency of 10 Hz to prevent aliasing.

The electrodes used were 10 mm Ag/AgCl (Med Associates TDE-022SN) and were applied with an electrode paste of 0.5% saline in a neutral base (Med Associates TD-246). The electrode surface, which was surrounded by a 2 mm high Teflon rim, was covered with a film of electrode paste even with the rim and the electrodes were fastened to the distal phalanxes of the first and second fingers of the non-dominant hand by a loop of adhesive skin tape. The electrode cables were secured to the palm by a third piece of tape to minimize any mechanical motion being transmitted to the electrodes.

Stimulus timing was derived from a pseudorandom number generator (Marsaglia, & Zaman, 1987), which was seeded from the system clock at the start of a session and in contrast to earlier studies (Radin, 1997a,b, Bierman & Radin, 1997, 1998), our participants were not required to initiate each trial; rather, the experimental session was "free running," and contained no stimulus-timing cues. The inter-stimulus interval was 60 ± 20 s and comprised the following sub-intervals: 5 s prestimulus period, 1 s stimulus duration, 24 s after stimulus time (i.e. a total of 25 s post stimulus according to the usual definition), 10 s fixed delay, and a random delay in the interval [0,40] s.[1]

Stimulus type was determined by electron shot noise within a true random number generator that was developed in the physics department at Ulm University in Germany.[2] This device, when hashed[3] with a pseudorandom generator passes the "Gold Standard" for random number generators—The Die Hard tests.[4] This generator was sampled at the end of the 5 s prestimulus period to determine the stimulus type—binary one indicated an audio stimulus, a binary zero indicated a silent control stimulus. In this way the decision as to the stimulus type occurred after all prestimulus data were recorded and stored in memory. This fact, coupled with the inherently indeterminate nature of a true random number generator of this type, constrains the interpretation of a positive outcome to the experiment. Such an interpretation must be based either upon chance, retrocausality, or upon psychokinetic effects acting upon the random number generator.

The audio startle stimuli were derived from computer-stimulated white noise of one second duration and 97-dB intensity and saved as a WAV file, while control stimuli comprised one second of zero signal, or silence. The participant wore sound isolating

[1] The 5 s prestimulus period only served as a time to transfer the skin conductance data from the serial port buffer to the computer memory. In the analyses, 3 s was used as the prestimulus period and was defined relative to the stimulus onset. During the prestimulus period, the data collection software was quiescent waiting for a 5-second timer to elapse.
[2] Details of this generator may be found at http://hlhp1.physik.uni-ulm.de/~freitag/spinoffs.html.
[3] The data from the RNG was exclusive or-ed with numbers from a pseudo-random number generator.
[4] See http://stat.fsu.edu/~geo/diehard.html for information and source code of these tests.

headphones[5] during the trial ensuring a low background sound level during the experiment sessions.

Before a session began, participant information such as name, age, and gender and session parameters such as number of stimuli and stimuli timing data were entered into screens provided by the data collection program.

Sessions consisted of 20 stimuli divided between audio and control, or silent, stimuli on a random basis. Thus, on the average, there were about 10 audio and 10 control stimuli in a session which lasted approximately 25 minutes, given that the inter-stimulus interval was 60 ± 20 s. The sequence of events during a session was a follows:

- The participants were given a brief overall project history and a complete description of the experiment. In addition, the participants were asked to review and sign an informed consent document.
- The participants were asked to rinse their hands in water and dry them thoroughly.
- An experimenter attached the electrodes as described above.
- The experimenter and participants watched the skin conductance on a display to demonstrate how their skin conductance changed if they moved or were startled by a clap of the experimenter's hands. They also conversed for about 10-15 minutes to allow the electrodes to equilibrate with the skin and to give the experimenter an opportunity to check that the equipment was performing normally.
- The participant was instructed how to abort the session, if necessary, by clicking on a button labeled "Abort." They were also told that at the completion of the run they would hear a message on the headphones saying that the run was complete. They were asked to then call out for the experimenter who would disconnect them and conduct an immediate analysis of the data.
- The participants were instructed to keep their eyes closed and to try to maintain a state of vigilance, or active expectation, during the trial time. They were also asked neither to employ any meditative techniques during the session time nor to fall asleep.
- The experimenter terminated the setup period by setting the system to automatic and initiating the session. At this point, the skin conductance plot shown on screen would go blank.
- The experimenter left the room for approximately 25 minutes and waited outside the door until the call from the participant at the end of the session.
- After the participant was disconnected from the monitoring equipment, the experimenter analyzed the data to provide feedback to the participant.

[5] Radio Shack # 33-1198

Data Collection Software

The data collection program was written in Microsoft's Visual Basic 6 and run on a PC under either the Windows 2000 or Windows XP Professional operating system. Great care was taken in the design of this code such that all steps in the computer program should be identical in the prestimulus region regardless of future stimulus type. Thus, even the computer was "blind" when collecting the skin conductance data, right up to the last point in the prestimulus region, as to what the next stimulus type would be.

At the start of the data collection period, the stimulus sounds, comprised of two identical size "WAV" sound files, were loaded from disc into memory.[6] Skin conductance data collection started when the experimenter initiated the automatic mode and left the room. At that time, a 135 ± 15 s cool down period timer was started automatically. When complete, the program went into a 20-fold loop and collected the experimental data in accordance with the stimulus timing described above.

After the final stimulus a further 30 s of skin conductance (SC) data were collected after which a message signaling the end of the trial was played to the participant's headphones. At that time, the skin conductance data was written to disk.

During the code's operation data acquisition was interrupt driven by the various timers and no data were stored to disk; thus, there was no code related disc activity. Occasional disk activity was observed to occur during test sessions, which resulted from the normal swapping activity of operating systems, and this did not interrupt the skin conductance data stream into the serial port buffer.

Participants

Participants for this experiment were drawn from the public via word-of-mouth and from various advertisements on the Internet. A lower age limit of 18 was pre-specified and participants who were discovered during conversation with the experimenter to be seriously ill or on medication were excluded. Since the protocol was not counterbalanced as to the number of audio versus control stimuli, we decided in advance to reject sessions with less than six stimuli of either type, to reduce the variance of the within session effect size.[7] The pilot series was specified to include 125 first-time participants, those being individuals who had not previously contributed to this experiment, and excluded the two experimenters.

Analysis

The analysis of the data was carried out in two different ways, each of which could be applied to a single session contributed by one participant or to multiple sessions contributed by a number of participants.

---

[6] The silent control WAV file consisted of all zeros but was the same size as the stimulus WAV file.
[7] A two-tailed binomial probability of exclusion of 0.041.

Analysis of Proportions of Skin Conductance Responses

Our primary analysis was to analyze the difference between the proportion of skin conductance responses prior to audio stimuli and prior to controls. This method was originally used by Vassy (1978) in an experiment based upon a classical conditioning paradigm.

To detect a non-specific skin conductance response, we used a two-fold cubic spline interpolation method to both smooth the skin conductance data and to assure, by definition of the cubic spline method, that the first derivative was well behaved.[8] Figure 1 illustrates the method used on a single skin conductance response from one of the datasets.

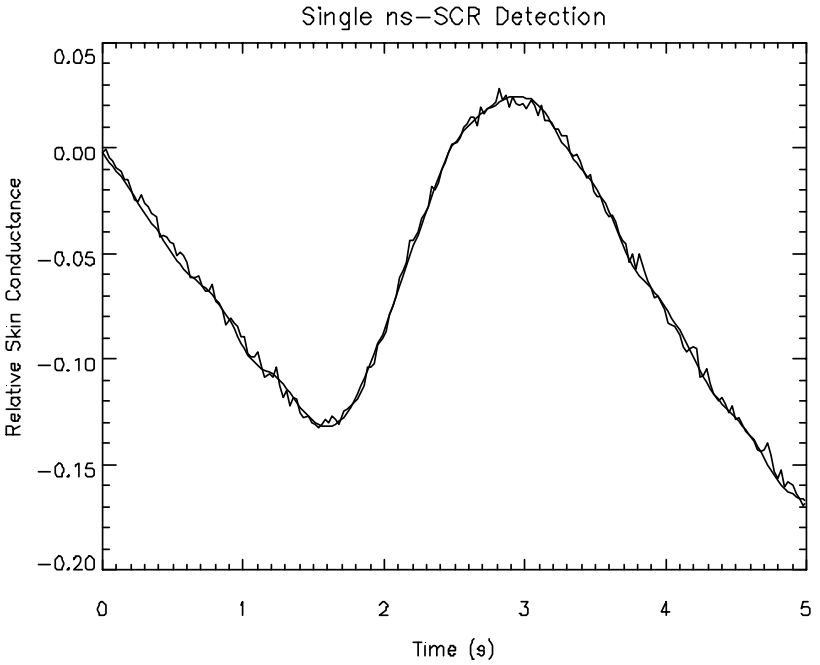

Figure 1. Spline Interpolation Method of Skin Conductance Response Detection

The "noisy" line in Figure 1 is a typical skin conductance response. (We have added a small amount of Gaussian noise to the original skin conductance data to illustrate the general method.) The solid smooth curve is the result of initially sampling the data at every eighth point, computing the cubic spline interpolation, and then using that spline to obtain a skin conductance approximation of the same length as the original record. Our definition of a skin conductance response required that a minimum be followed by a maximum in the prestimulus region. In this example, the algorithm computed the amplitude of this skin conductance response to be 0.156 μSiemens. Additionally, to qualify as an skin conductance response in our prestimulus region of three seconds, we required that the amplitude, computed from the above algorithm, must have exceeded a

[8] By two-fold, we mean we first down-sample the raw data as a smoothing procedure, and then up-sample the spline to the original data length.

specified threshold, and that only a single skin conductance response be present in the region. Because of the slow response times of skin conductance changes, it was rare that two skin conductance responses were found in the prestimulus region, and we ignored these epochs. To define the detection threshold, we computed the mean amplitude of all the skin conductance responses in a single session, removing the data from [-6,10] s around each stimulus. The skin conductance response amplitudes in the remaining regions were found by applying the same algorithm as above. However, we required that the amplitude of the skin conductance response be greater than or equal to 0.005 μSiemens to exclude instrumental noise in very stabile participants and low amplitude pulses such as those due to heart pulse, as was observed in a few participants. To account for differing skin conductance levels between participants, we defined the threshold for a skin conductance response for each session. Specifically, we required that a given skin conductance response meet or exceed 0.1 times this mean skin conductance response amplitude for each participant. This value of 0.1 was arbitrary, and as we will show below, our results are insensitive to this parameter.

We defined the lability of a participant as the number of skin conductance responses that exceeded this threshold divided by the total number of seconds in the above analysis period. Participants at the upper end of the lability range (e.g., 0.1 responses per second) are conventionally denoted "labile," while those at the lower end are called "stabile."

Under the null hypothesis, the proportion of skin conductance responses prior to an audio stimulus should be equal to that before a control, silent "stimulus." The analysis proceeded as follows.

Consider two types of intervals: (1) a prestimulus period just prior to sound and (2) a prestimulus period just prior to a silent control. We counted a skin conductance response as having occurred in either of these intervals if the criteria that were described above were met:

The proportions prior to audio stimuli, p_s, and prior to control, p_c, stimuli are defined as:

$$p_s = \frac{\#\,of\ skin\,conductance\,responses\ prior\ to\ audio\ stimuli}{\#\,audio\ stimuli = N_s},$$

$$p_c = \frac{\#\,of\ skin\,conductance\,responses\ \ prior\ to\ control\ stimuli}{\#\,control\ stimuli = N_c}.$$

The standard relation for the Z-score for the difference between two proportions is given by (Larsen & Marx, 1986):

$$Z = \frac{p_s - p_c}{SD},$$

where the standard deviation, SD, is given by:

$$SD = \sqrt{p(1-p) \times \left[\frac{1}{N_s} + \frac{1}{N_c} \right]},$$

and p is given by:

$$p = \frac{\left(p_s N_s + p_c N_c \right)}{N_s + N_c}.$$

The analysis was the same for a single session as for data combined across sessions. In the latter case, the proportions and stimulus counts reflected the combined data.

Epoch Analysis

Our secondary approach was an extension of the ensemble-averaged skin conductance levels method, which we used so that our results might be compared to earlier studies. Often what is done in the analysis of psychophysiology data is to detrend data epochs prior to ensemble averaging. However, we modified the usual approach to leave open the possibility that the skin conductance prestimulus response might consist of a gradual change in slope in the skin conductance data prior to audio stimuli as compared to controls. A detrend of the epoch of interest would have eliminated such a change and rendered it unobservable. Instead, to capture the possibility of psi-mediated trends in the region of interest, we fitted the skin conductance data in a two-second window prior to the start of our three-second prestimulus region with the two-fold cubic spline interpolation as described above. From the spline fit, we obtained the slope of the data 25 ms, or one data sample, prior to the start of the prestimulus region and used it to detrend the prestimulus data.[9] In this way the three second prestimulus region was detrended while preserving the possibility of detecting a change of slope in this region should it occur. Figure 2 shows an example of such a calculation.

The smooth curve in the interval [-5,-3) s is the result of the two-fold spline interpolation described above. The extended straight line in the [-3,1) s region is the extension of the slope computed from the spline at –3.025 s (i.e., one data sample prior to the start of the prestimulus region). The middle dashed curve in this region is the raw data clamped at –3 s for convenience. The top dotted curve is the residual; that is, the raw data minus the trend line. It was this residual that was used for the following ensemble-averaged epoch analysis. We note that in this example, the data contains a skin conductance response in the prestimulus region and that the raw data in the [-5,-3) s region is overlaid by the spline interpolation.

[9] By definition, a spline is a continuous function whose first derivative is defined at every point. Thus we can compute the slope at the single last data sample.

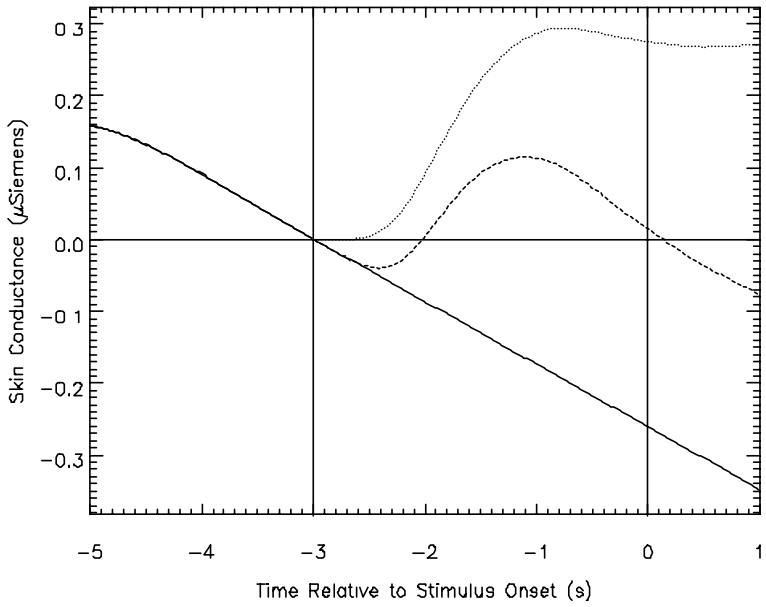

Figure 2. Example of Trend Removal

Filters

To account for possible outliers that might have affected the ensemble-averaged epoch analysis, we applied two types of filters to each epoch. Later we will discuss the results with no such filtering. Another way to think of these filters is to consider them as two asymmetric thresholds. If the data exceed either, the stimulus is rejected.

Slope Filter

For each of the prestimulus epochs, the analysis shown in Figure 2 yields a slope that we used to detrend the data. Figure 3 shows the distribution of these slopes for all of the 2,500 stimuli in the pilot study regardless of stimulus type.

First, we note that the mode of the distribution is slightly negative indicating a general tendency for skin conductance to be decreasing at the 2,500 points where the slopes were calculated. This reflects the overall decrease in skin conductance observed in most sessions. However, we also observed that there were a number of large amplitude slopes in the tails of the distribution. We arbitrarily set slope limits corresponding to the upper and lower 2.5%—the vertical lines in Figure 3.[10] In the epoch analysis, we rejected epochs that fell outside these limits.

[10]We will return to the point of how sensitive the analysis is to this threshold below.

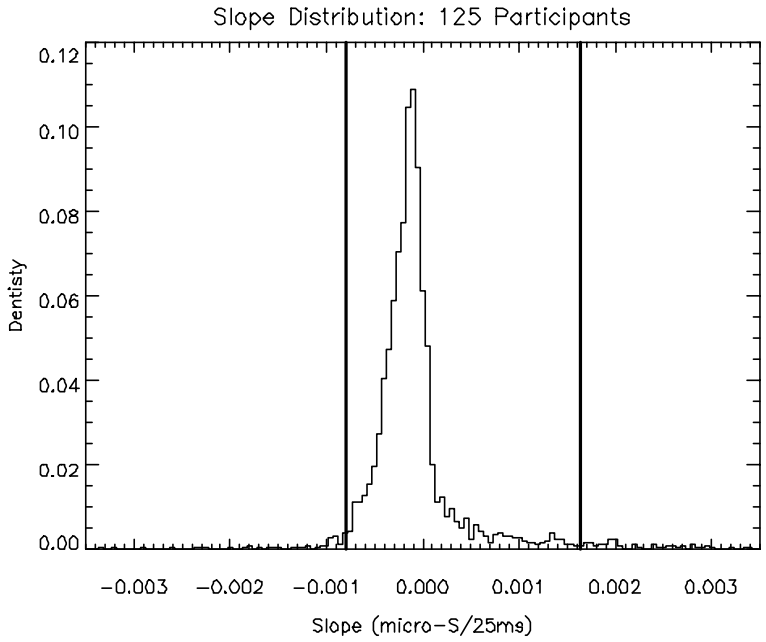

Figure 3. Slope Distribution for all Stimuli in the Pilot Study

Transition Filter

To account for other outliers, we constructed a transition filter as follows. We considered the raw skin conductance data in the prestimulus region starting at three seconds prior to an audio or control stimulus. We computed the fractional change of skin conductance in these intervals as:

$$\Delta = \frac{\left| SC_{\max} - SC_{\min} \right|}{SC_{ave}}.$$

The sign of Δ was determined by the position of the maximum skin conductance relative to the minimum SC. If the maximum occurred before the minimum, Δ was defined as negative. Figure 4 shows the distribution of Δ for all prestimulus regions in the pilot study.

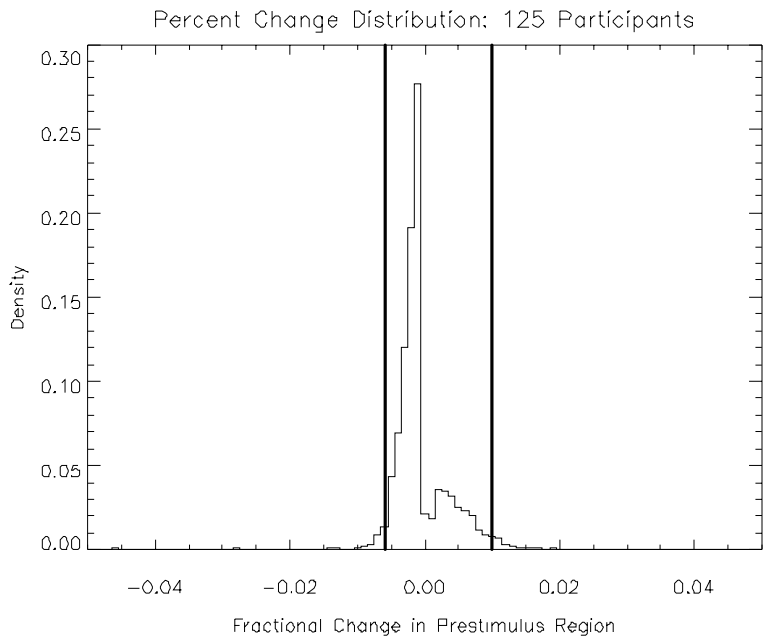

Figure 4. Fractional Skin Conductance Transition Distribution

As in the case of the slope filter, we arbitrarily rejected epochs that produced transitions in either of the 2.5% tails.

Ensemble Averages

We calculated the ensemble averages of the prestimulus regions of [-3,0) of the residuals that were derived from the detrending described above for those epochs that survived the two filters for the combined dataset of 125 individual sessions. These ensemble averages were computed for pre-audio and pre-control epochs. We also calculated the prestimulus area difference (PAD) as the area under the pre-audio average minus the area under the pre-control average. The PAD was the independent variable for this analysis.

We used a standard randomized permutation analysis (RPA) technique to determine the degree to which the PAD was statistically different from zero, the chance expectation value. In a single RPA pass, we randomly permuted the assignment of data epochs to audio or control stimuli, keeping the same number of epochs in each group and computed new ensemble averages and a new value of PAD. We conducted 20,000 such passes to produce a distribution of PAD's, which, because of the central limit theorem, tends rapidly to become normal. We computed the observed p-value in the usual way from this distribution.

Results

The primary analysis was the difference between the proportions of skin conductance responses prior to audio and control stimuli. The secondary analysis was the ensemble-averaged epoch analysis. The 125 participants comprised 60 males and 65 females

ranging in age from 20 to 74 years (median age 46 years). Labilities ranged from 0.0012 to 0.2468 skin conductance responses per second.

Analysis of Proportions of Skin Conductance Responses

As described above, we computed the proportion of skin conductance responses prior to audio and control stimuli across all participants. We also used the expression for the Z-score shown above to assess the statistical significance of the difference between these two proportions.

Table 1. skin conductance response Results for N=125 First Time Participants

Interval Type	Number of skin conductance responses	Number of Stimuli
Before a Control	56	1181
Before a Stimulus	105	1319

We computed a *Z*-Score of 3.27 and a per stimulus effect size of 0.0901 ± 0.0275 for a *p*-value of 5.4×10^{-4} (1-tailed). On a per participant basis we compute an effect size of 0.292 ± 0.089[11].

Epoch Analysis

Figure 5 shows the results of the separate ensemble averaging prior to audio and control stimuli across all 125 sessions. We show the 3 s prestimulus region and only 1.5 s of post stimulus to avoid compressing the pre-stimulus region of interest by the large post stimulus response, which is approximately 100 times larger than the y-axis scale. The dashed lower curve is the average prior to the 1102 silent controls (i.e., those epochs that survived the filters) and the upper solid curve is the average of the 1205 filter-surviving epochs prior to an audio stimulus. The error bars are one standard error of the averages at the time shown. The prestimulus area difference (*PAD*) is the area between the two averages and is 0.00873 μSiemens·s.[12]

To assess the significance level, we used the RPA approach described previously. The fraction of the resulting array of 20,000 areas that was equal to or greater than the observed area constituted the p-value. By using an inverse normal transform, we computed the *Z*-score and effect size. Figure 6 shows the distribution of areas for the data shown in Figure 5. The histogram is the distribution of areas and the smooth curve is the best-fit normal curve to the histogram. The vertical line at about

[11] Including the four trials which were rejected as having less than six stimuli of one type would have increased the quoted effect sizes.
[12] The apparent SCR beginning at about 300 ms post stimulus is a muscular flinch response.

PAD = 0.009, is the observed value in the analysis leading to a *Z*-score of 3.16 and an effect size per stimulus of 0.066 ± 0.028. We observed a null hypothesis mean (i.e., based upon the randomised permutation methid) of 0.00003 μSiemens·s, which is close to the expected PAD value of zero under the null hypothesis.

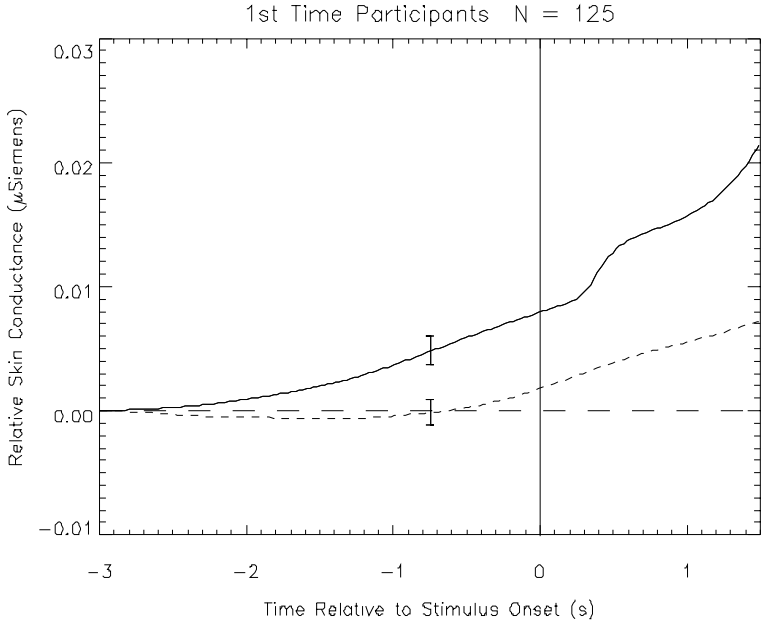

Figure 5. Epoch Analysis of 125 First Time Participants

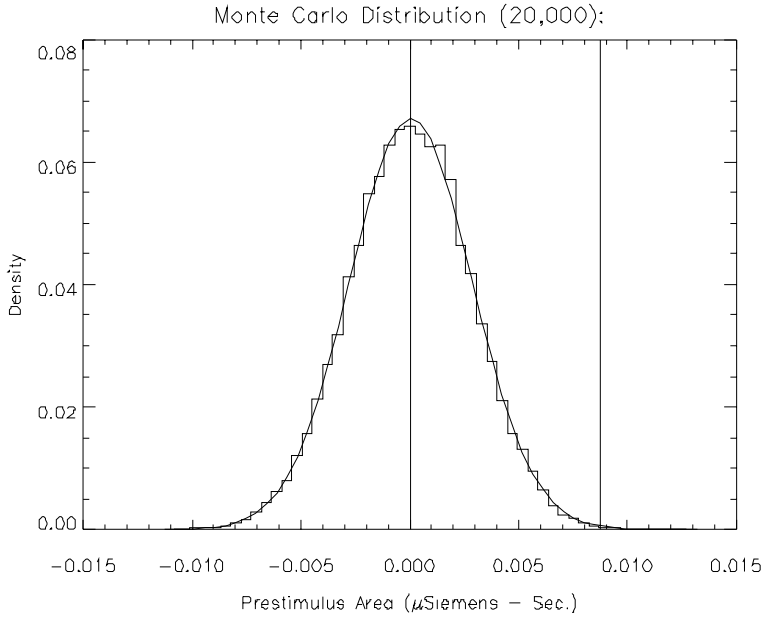

Figure 6. Randomized Permutation Result

Potential Artifacts

There are a number of sources of potential artifacts for both prestimulus response data analyses.

- Cueing. Subtle cues allowed the participant to know the up-coming stimulus type.
- Expectation. Either the rate of skin conductance responses or the skin conductance level could have increased due to stimulus hunger.
- Stimulus Generator. There were non-random anomalies in the stimulus generator that allowed a participant to infer the next stimulus' type.
- Programming Errors. The observed effect arose because of errors in either the data collection or analysis code.
- Data Collection Computer Anomaly. The observed effect was due to an unknown mechanism in the data collection computer, which caused it to record skin conductance responses in a biased manner.
- Participant or Experimenter Fraud. The results occurred because of cheating.

We address each of these points below.

Cueing

Both the experimenter and the participant were blind to the up-coming stimulus choice. The data collection code was also "blind" as well, since the stimulus type was determined by a hardware RNG 25 ms *after* the last skin conductance data point in the prestimulus region had been saved in computer memory. The data collection computer ran identical code for each of the two possible future stimuli types up to the moment of stimulus presentation. In addition, the audio files that were played as the stimuli were resident in memory throughout the run and thus there were no code-related disk accesses. Cueing could not therefore contribute to the observed outcome.

Expectation: Skin Conductance Response Analysis

Lately there has been theoretical interest in possible contributions to apparent prestimulus response by expectation effects (Dalkvist et. al, 2002. Wackerman, 2002). The approach is to assume that there is a monotonic increase in the dependent variable as a function of time between adjacent arousing stimuli; simulate this in a computer program; and then to compute the resulting Z-score reflecting this assumption. The discussion has focused upon skin conduction level shifts, but the concept could apply to our dependent variable as well if there were an increase in the rate of skin conductance responses as time elapses between adjacent audio stimuli. However, these studies also show that in analyses in which data is summed across subjects, as here, the effect can only be very small and this is confirmed by the two tests which follow.

In the first test we moved the analysis period back 3 s to [-6,-3) s relative to the audio and control stimuli and recalculated the skin conductance response statistic. By repeating this process of moving the whole analysis backwards, we computed the effect size as a

function of prestimulus intervals before the stimulus. Figure 7 shows the result including the effect size for the prestimulus interval of [-3,0).

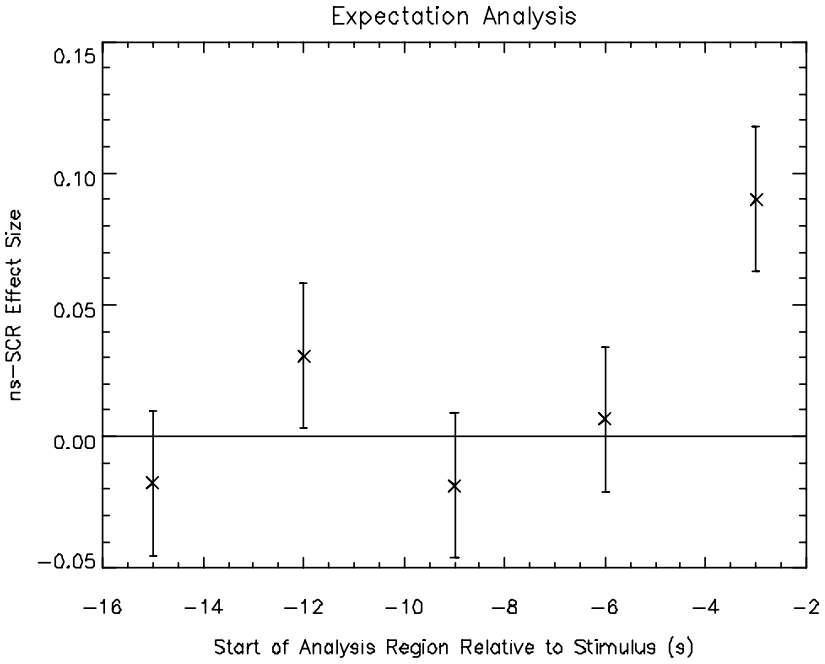

Figure 7. Proportion Z-Score for Audio Stimuli

The effect size in the 3 s window immediately prior to the prestimulus period is significantly smaller than that in the prestimulus region (i.e., t_{dff} = 2.15, df = 2636, p = 0.016) and the effect size remains at mean chance expectation for an additional three earlier intervals.

We now have the following three conditions: An excess of skin conductance responses in the prestimulus epoch prior to audio stimuli, no such excess over chance expectation in the preceding epochs and a randomized inter-stimulus interval of 40 s to 80 s. Any expectation effect must appear as an increase in the dependent variable (i.e. rate of production of skin conductance responses), which is a monotonically increasing function of the time since the last arousing stimulus. Such an effect, were it to occur, could not give rise to an increase in the skin conductance response rate solely in the epoch prior to the next startle stimulus, *a fortiori* when the timing of that stimulus was randomly varied and unknown to the participant.

A second approach to the expectation issue is, perhaps, more direct. By definition, expectation artifacts of any kind should contribute more the longer the time since the last audio stimulus. The mean of the distribution of times since the last audio stimulus was 106.886 s for audio stimuli and 109.482 s for controls (t_{diff} = -0.838, df = 2498, p = 0.800) and the distributions of these times were not significantly different (Kolmogorov-Smirnov p = 0.232). Since these two sets of times are statistically indistinguishable, the expectation bias model cannot produce differences in skin

conductance response rates. We confirmed this in our data since we found a Pearson's *r* correlation of 0.056 ($t = 0.448$, $df = 64$, $p = 0.382$) for effect sizes correlated against the mean time between adjacent audio stimuli.[13]

This result in itself is sufficient to reject an expectation bias hypothesis; however, coupled with the result that the effect size was at chance in the [-6,-3) interval (and earlier ones—see Figure 7) and is significantly lower than that for the actual prestimulus interval of [-3,0), we conclude that expectation did not account for the proportional skin conductance response rate in this study.

Stimulus Generator

As we have pointed out above, the hardware stimulus generator and its associated software passes the "gold standard" test for randomness; however, it remains possible that some momentary stimulus-to-stimulus dependence could be sensed by a participant. To examine this possibility we computed the auto-correlation for the actual stimulus sequence in the study for ± 20 lags. If there were no serial dependencies between the given stimulus and up to 20 previous or future stimuli, then the auto-correlation function would be unity at zero lag and statistically zero for all other lags. In ± 20 lags, we would expect one significant autocorrelation. We found none. Thus, there were no simple inter-bit patterns in the stimulus sequence and therefore a participant would be unable to anticipate correctly the up-coming next stimulus.

Programming Errors

The data collection code was shared between the two experimenters (May & Spottiswoode). To guard against both a conceptual error and a possible coding error in this shared code, we replaced the human participant with a resistor that represented a typical skin conductance level (i.e., 220 kΩ) that was placed between the earphones to crudely simulate a human. We would expect no differences between the prestimulus regions regardless of stimulus condition. Not only did we not observe any differences, but when we also measured the skin conductance of participants without any overt stimuli, we also saw no significant effects and no post-stimulus responses.

To guard against coding errors in the analysis code, May & Spottiswoode conceptually agreed on the analysis method. They then both coded that analysis with different approaches using different computer languages and sharing no code. The results were not accepted unless the calculations agreed to within machine precision for floating point calculations. This method guards against coding errors but does not eliminate possible conceptual errors.

[13] To keep from possibly skewing the result we only used data that contained skin conductance responses in the prestimulus region.

Conceptual errors were also very unlikely in that the primary analysis was a simple skin conductance response counting exercise. Therefore, programming errors did not contribute to the positive results of the study.

Data Collection Computer Anomaly

As an additional check for unknown artifact mechanisms, we constructed a pseudo participant as a control condition for the entire experiment. A possible skeptical response to these results might be that some effect occurring in the data collection computer acted so that the recording of the large post-stimulus signals caused a change in the pre-stimulus data recorded slightly earlier. Although it is difficult to imagine how this could occur since modifications of a computer's memory by an external source usually result in a computer crash, we thought that it was incumbent on proponents of novel data, such as this, to explore even remotely unlikely artifacts. More generally, we wanted to investigate what would be the result of replacing the human participant with a system that would interact with the data collection hardware and software in a way similar to a real participant.

Referring to Figure 8, a sound pressure level meter[14] was placed between the earpieces of the headphones used by the participant in the prestimulus response experiment. The output signal representing the sound level was passed through a 16-bit analogue to digital converter[15] and then to a threshold detector, which was set to trigger when the sound level at the headphones exceeded approximately 85 dB. This trigger signal was then delayed for 2 s (the mean time interval between the onset of the audio stimuli and the start of the post-stimulus response in humans). This delayed trigger was then passed to a routine that generated a simulated post stimulus response signal. This was comprised of a sigmoid leading edge rising to a maximum at approximately 3 s after triggering followed by a slow exponential, decay with a time constant of 10 s. Also running in the simulation software were eight skin conductance response generators, exactly like the post-stimulus response generator in design, which were set to have a fixed probability of firing per second and a random amplitude between zero and an adjustable maximum.

[14] Radio Shack Digital Sound Level Meter #33-2055

[15] The A/D converter was an HP/Agilent E1326B and the D/A an HP/Agilent 1328A both running in an HP/Agilent E1301B VXI chassis connected by GPIB interface to a PC running Windows 2000 and Labview

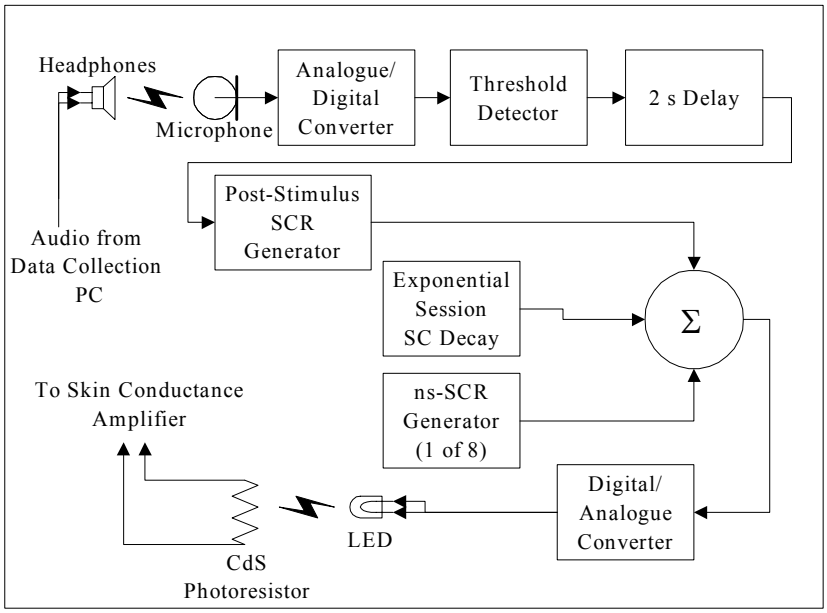

Figure 8. Simplified Block Diagram of the Participant Simulator

Multiple skin conductance response generators were used to allow for more than one skin conductance response to occur at once given that each one could not be used again until after many seconds had elapsed and it had reached zero output. To simulate the gradual decline in skin conductance level during an experimental run seen in most participants, a signal with a very slow exponential decline and a settable floor was also generated. All these signals were added and their sum passed to a digital-to-analogue converter, which drove an LED. This LED was placed next to a Cadmium Sulphide (CdS) photo-resistive cell, both being shielded in a light-tight enclosure. The CdS cell was connected to the electrodes, which would otherwise be attached to the participant. Thus, the same skin conductance monitor that was used in the prestimulus response experiment saw a varying conductance in the CdS cell, which closely mimicked the response of an actual participant. Note also that in this arrangement, there was no electrical connection between the simulation computer and hardware and the data collection hardware used in the prestimulus response experiment, the links between the two systems being acoustic and optical.

The parameters were set such that simulated data had typical tonic skin conductance levels (2 – 20 μSiemens) and post stimulus SCR amplitudes (1 – 3 μSiemens). By setting probabilities amplitudes that matched real data for firing and maximum amplitude of the skin conductance response generators, this simulator generated skin conductance data that were very similar to those from a real experimental session. Stabile and labile participants were simulated by varying these parameters.

The simulator was run against the data collection computer using exactly the same data collection software and session parameters as the 125 experimental sessions for 125 pseudo-sessions. Three settings for lability were used each for approximately a third of

the simulated sessions. The results are shown in Figure 9 using the appropriate ± 2.5% filter thresholds for this data set.

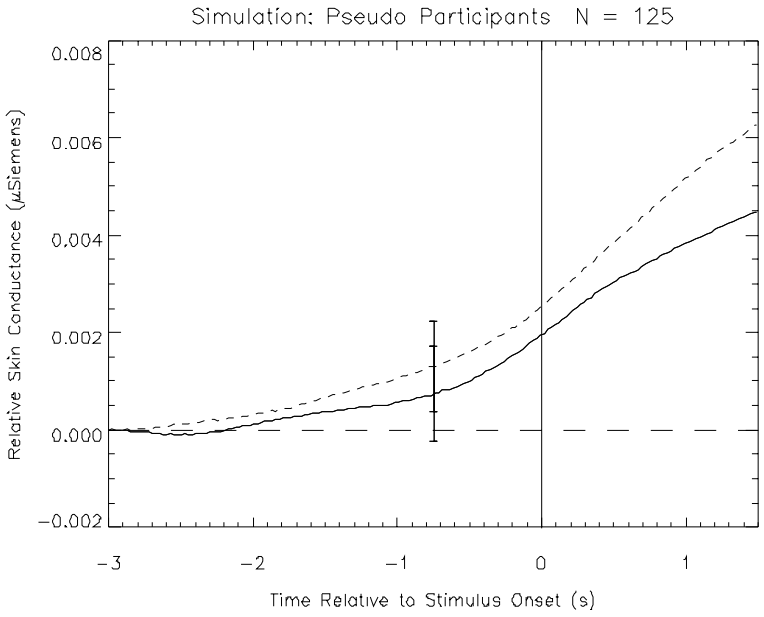

Figure 9. Epoch Analysis Results from Simulation

The dashed curve in Figure 9 represents the epoch-averaged simulated skin conductance prior to silent controls whereas the solid curve represents the average prior to audio stimuli. The prestimulus area difference in the simulation was –0.001 (p = 0.663, ES_{epoch} = -0.009 ± 0.0283), and the number of skin conductance responses prior to audio and silent controls was 104 and 95 (p = .249, $ES_{proportion}$ = 0.0191 ± 0.0283), respectively. That is, both effect sizes were within one standard error of zero. Therefore, we concluded that no artifact-inducing effects occurred in the data collection hardware and software. In fact, the entire processing sequence from data collection to final analysis was tested in these simulation results, which demonstrate that replacing the human participant with a near simulacrum in terms of skin conductance behavior caused the observed prestimulus response effect to vanish.

Fraud

We examined two possible types of fraud: by participant or by experimenter.

Participant Fraud

For participant fraud to occur, an individual, who was not aware of the technical details of the experiment, had a maximum of 28 minutes to replace the actual data file with a clandestine pre-defined, fraudulent one or to try to manipulate the data collection system or generate signals by motor movements to mimic the prestimulus response effect.

To eliminate the possibility of participant fraud, all programs and data were 128-bit encrypted on the computer's disk. In addition, the data file structure was confidential and

not shared with anyone. During the time of the experiment, the data collection computer was not connected to the Internet. Therefore, the first mechanism is extremely improbable. Fraud based on manipulation of the data collection system or generation of signals by movements is moot since the stimulus type, which would have to be known in advance to generate false signals, was unknowable until their occurrence.

Therefore, the results of this study cannot be accounted for by participant fraud.

Experimenter Fraud

The only real protection against experimenter fraud is independent successful replication. However, at a minimum, experimenter fraud in this experiment would imply a conspiracy of two. In fact, the effect sizes for the skin conductance response analysis were not statistically different between May and Spottiswoode (t = 0.616, df = 1317, p = 0.270; ES_{May} = 0.0711 ± 0.0411, N = 56 sessions, $ES_{Spottiswoode}$ = 0.1052 ± 0.0371, N = 69 sessions), indicating either a matched fraud or a real effect. The authors would be pleased to assist other experimenters with replication attempts, which is the only definitive way to resolve this explanation for the results.

Discussion

We have observed a significant prestimulus response in first time participants in this pilot study using the skin conductance response proportion measure. In the language of psychophysiological work on skin conductance, we would call this the phasic component of prestimulus response. In contrast, the epoch analysis method is also sensitive to the tonic component in the skin conductance. We now consider a difficulty with the ensemble-averaged epoch analysis.

The significant (i.e., Z = 3.12) prestimulus area difference shown in Figure 5 above appears quite impressive. The residual skin conductions prior to controls are statistically indistinguishable from zero, while the epoch average for the pre-audio region shows a significant positive trend.

To illustrate the problem, we began by re-analyzing the same data in the manner of other studies (Bierman & Radin, 1997; Radin, 1997b), that is, by clamping the data at the start of the prestimulus region and using no detrending or epoch filtering, and a non-parametric RPA method to assess statistical significance. Figure 10 shows the result; the solid curve is the average prior to audio stimuli and the dashed curve is the average prior to silent controls.

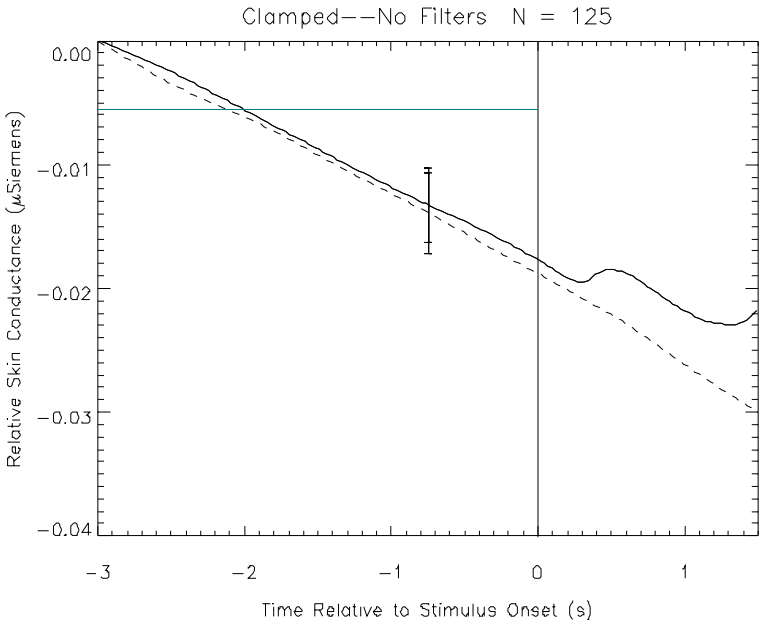

Figure 10. Data Clamped at –3 s with no Filters or Detrending

The RPA *Z*-score has suffered a significant drop (Z_{diff} = 2.06) from 3.16 to 0.24. Yet, using the proportion statistic, this same data contains a significant effect. Thus, we emphasize here that a significant result from the usual sort of epoch analysis and RPA assessment is not necessary to observe significant prestimulus response. We now consider whether epoch averaging is sufficient to indicate prestimulus response.

We begin by assuming that a prestimulus response contains two components: a differential level shift, the tonic component and a differential skin conductance response count, the phasic component. We may examine the tonic component alone by removing those sessions that produced a skin conductance response in the prestimulus region regardless of stimulus type. We lowered the skin conductance response detection threshold to the minimum noise level of 0.005 μSiemens to account for even smaller phasic contributions than allowed in our skin conductance response analysis above. We found that 49 sessions met this criterion.

Analyzing these data as before including the filtering, we found that removing most, if not all, of the phasic contribution reduced the RPA-derived effect size from 0.0657 ± 0.0275 to 0.0281 ± 0.0445 (t_{diff} = 0.718, df = 1821, p = 0.236). Still, there remained a slight increase of skin conductance prior to audio stimuli compared to controls. This might arise because of skin conductance responses of lower amplitude than 0.005 μSiemens or because of threshold effects in the filtering. To examine the latter case, we analyzed the same set of data clamping the data at –3 s and removing all filters.

We found that any apparent prestimulus response had vanished (ES = 0.0019 ± 0.0445, p = 0.476). Clearly, the threshold filters have some effect upon the observation of prestimulus response though epoch analysis. We then examined the effects of filter

parameters by varying the thresholds for both filters from 0 to 5 percent in steps of 0.05 percent. The results comprise effect sizes, *Z*-scores and rejected epochs for 121 possible combinations of the two filter thresholds. An indication of the results is shown in Table 2.

Table 2. Filter Threshold Analysis Sensitivity

	Threshold (±%)		
	2	2.5	3
Total Epochs	912	908	869
Z-Score	1.555	0.837	-0.052

We see that a 0.77% increase of analyzed epochs (7 more than 905) increases the *Z*-score by 85% (from 0.84 to 1.55). Rejecting an additional 9 epochs, the *Z*-score goes from 0.84 to less than zero. Thus, we observe an extreme sensitivity to skin conductance response outliers, which we might expect with any sort of threshold analysis. It is important to note, however, a few large outliers just above or below the filter thresholds can drastically drive the resulting RPA-derived Z-score. This demonstrates that the non-parametric partial permutation method does not solve possible outlier problems.

With regard to simple clamped epoch analysis, which contains no filters and thus no thresholds, we conclude that a reported significant difference between arousing and calming (or control) stimuli *may* arise because of the phasic contribution. However, it also may arise because of a fortuitous amplitude distribution of a few outliers even with no thresholds to consider and randomized permutation analysis.

Thus, we urge caution in interpreting the results of such analyses, and suggest that an epoch analysis be conducted using medians rather than averages to examine the effects of outliers.

The differential skin conductance response analysis, however, was not dependent upon filter parameters. Our definition of a valid skin conductance response contained one free parameter: namely, the fraction of the mean skin conductance responses for the given session. We arbitrarily chose 0.1. Figure 11 shows the effect size for the differential skin conductance response count as a function of this fraction.

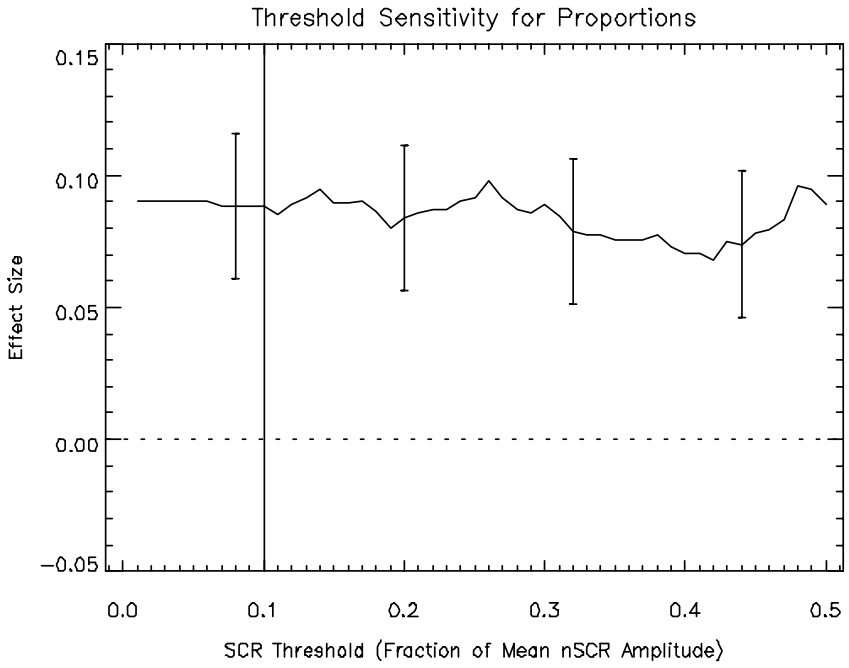

Figure 11. Effect Size Sensitivity to Threshold Parameter

The error bars are for one standard error for the effect size and the vertical line (red) at 0.1 indicated the value used in the analysis.[16] Therefore, our result is insensitive to the choice of discrimination threshold for a wide range of values.

Post Hoc Observation

We noticed that the number of audio (1319) and control (1181) stimuli in this experiment deviates significantly from chance expectation (Z = 2.76, p = 0.006 2-tailed, ES = 0.0552). While there is dispute as to the mechanism of these effects, the observable in the literature is that the effect size scales as $1/\sqrt{n}$ where n is the number of bits in the sequence (May et al., 1995). Using a summary of 12 years of random number generator experiments (Jahn et al, 1997, page 350) as an estimate of the "true" psychokinetic effect size, we computed an effect size per bit of 2.6×10^{-4} for 838,800 trials at 200 bits each. This corresponds to an effect size per bit for our study of 2,500 bits of 7.4×10^{-5}. However, our observed effect size per bit for the stimulus count differential is 0.0552. Since this effect size is 746 times larger than we computed for our 2,500-bit case based on the literature, it is unlikely that RNG PK[17] can account for our observed deviation. While it remains possible that RNG PK effect is strong at very low bit rates such as about one per minute, the data at a few bits per second do not show this trend (May, et al., 1995). Thus, our significant stimulus differential count remains a mystery.

---

[16] The one standard error is defined here is 1.0 / $\sqrt{number\ of\ audio\ stimuli}$ which is independent of threshold.

[17] Random Number Generator Psycho-Kinesis (RNG PK) is a protocol in which participants attempt to bias the output of a random source by means of their intention.

Conclusion

We have conducted a pilot prestimulus response study incorporating audio startle stimuli and silent controls in order to minimize inter-subject variance. We have observed a significant skin conductance prestimulus response as measured by a differential count of prestimulus skin conductance response counts. Explanations of this result based on cueing, expectation effects, hardware anomalies and fraud were examined and rejected as either impractical or inconsistent with the data.

We have compelling evidence that, at least in this data set, that prestimulus response occurs through phasic changes in skin conductance and not through slow tonic level shifts.

Acknowledgements

We thank Zoltán Vassy for inspiring the approach in this experiment and for many helpful discussions. We thank the Institute of Noetic Sciences for helping to recruit participants. We thank Diane Bassett for her tireless effort in recruitment as well. We also thank the Fundação Bial, the Institute of Noetic Sciences and the Samueli Institute for Information Biology whose financial support made this study possible.

References

Andreassi, J. L. (1989). *Psychophysiology: Human Behavior and Psychological Response*. Hillsdale, New Jersey: Erlbaum Associates.

Bierman, D. J. & Radin, D. I. (1997). Anomalous Anticipatory Response on Randomized Future Conditions. *Perceptual and Motor Skills,* **84**, 689-690.

Bierman, D. J. & Radin, D. I. (1998). Conscious and Anomalous Non-Conscious Emotional Processes: A Reversal of the Arrow of Time? In: *Toward a Science of Consciousness, Tucson III.* MIT Press, 1999, 367-386.

Bierman, D. J. (2000). Anomalous Baseline Effects in Mainstream Emotion Research Using Psychophysiological Variables. *Proceedings of Presented Papers: The 43rd Annual Convention of the Parapsychological Association,* 34-47.

Bierman, D. J, & Scholte, H. S. (2002). Anomalous Anticipatory Brain Activation Preceding Exposure of Emotional and Neutral Pictures. *Proceedings of Presented Papers: The 45th Annual Convention of the Parapsychological Association,* 25-36.

Bouscein, W. (1992). Electrodermal Activity, New York: Plenum Press.

Braud, W. & Schlitz, M. (1983). Psychokinetic Influence on Electrodermal Activity. *The Journal of Parapsychology,* **47**(2), 95-120.

Crithchley, H. D., Elliott, R., Mathias, C., J., and Dolan, R. J. (2000). Neural Activity Relating to Generation and Representation of Galvanic Skin Conductance Responses: A Functional Magnetic Resonance Imaging Study. *The Journal of Neuroscience*. **20**(8), 3033-3040.

Dalkvist, J., Westerlund, J. & Bierman, D. J. (2002). A Computational Expectation Bias as Revealed by Simulations of Presentiment Experiments. *Proceedings of Presented Papers: The 45rd Annual Convention of the Parapsychological Association,* 62-79.

Lang, P. J., Bradley, M. M. & Cuthbert, B. N. (1999). *International affective picture system (IAPS): Instruction manual and affective ratings*. Technical Report A-4, The Center for Research in Psychophysiology, University of Florida.

Larsen, R. J. & Marx, M. L. (1986). An Introduction of Mathematical Statistics and its Applications. Prentice-Hall, Englewood Cliffs, New Jersey, pp. 378-380.

Marsaglia, G., and Zaman, A. (1987). Toward a Universal Random Number Generator. Florida State University Report: FSU-SCRI-87-50.

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

Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H. and Bradish, G. J. (1997). Correlations fo Random Binary Sequences With Pre-Stated Operators Intention: A Review of a 12-Year Program. *Journal of Scientific Exploration*, **11**(3), 345-368.

Parkhomtchouk, D. V., Kotake, J., Zhang, T, Chen, W, Kokubo, H, and Yamamoto, M. (2002). An Attempt to Reproduce the Presentiment EDA Response. *Journal of International Society of Life Information Science*, **20**(1), 190-194.

Radin, D. I. (1997a). *The Conscious Universe*, Harper Edge, Harper Collins, New York, NY.

Radin, D. I. (1997b). Unconscious Perception of Future Emotions: An Experiment in *Presentiment. Journal of Scientific Exploration*, **11**(2), 163-180.

Vassy, Z. (1978). Method for Measuring the Probability of 1 Bit Extrasensory Information Transfer Between Living Organisms. *Journal of Parapsychology*, **42**, 158-160.

Wackermann, J. (2002). On cumulative effects and averaging artefacts in randomised S-R experimental designs. *Proceedings of Presented Papers: The 45th Annual Convention of the Parapsychological Association*, 293-305.

Acknowledgments

We thank Zoltán Vassy for inspiring the approach in this experiment and for many helpful discussions. We thank the Institute of Noetic Sciences for helping to recruit participants. We thank Diane Bassett for her tireless efforts in recruitment as well. We also thank the Fundação Bial, the Institute of Noetic Sciences and the Samuelli Institute for Information Biology whose financial support made this study possible.