KPU Registry ID Number: 1005 Date revised, 5th December 2014
(This revised version was submitted prior to any data collection)

Study Registration

For the Koestler Parapsychology Unit Study Registry

1. The title or name of the experiment (for listing the experiment in the registry).

Single-trial presentiment experiment.

2. The name, affiliation, and email address for the lead experimenter(s) for the study.

Julia Mossbridge, Northwestern University, j-mossbridge@northwestern.edu

3. A short description or abstract of the purpose and design of the experiment.

The motivating analysis and the pilot experiment were both performed before I knew of this registry. The motivating analysis is reported in the Mossbridge et al. paper submitted to the 2011 Parapsychology Association meeting. In an attempt to avoid potential expectation bias problems, the analysis was performed only on physiological data from the first trial of a multiple-trial guessing task. The analysis showed that on the first trial, skin conductance changes in the 10-s period preceding feedback as to the correctness of the guess were in significantly different directions for men and women (men's skin conductance went higher on correct guesses than incorrect guesses, for women the trend was the opposite direction), and skin conductance changes preceding correct first trials and incorrect first trials were significantly different in men (but not women). There were no significant effects for heart rate.

We performed a subsequent pilot experiment in which participants performed single trials of each of two presentiment conditions, in an attempt to determine which would be most effective in producing a presentiment response in a single trial. In both conditions, heart rate and skin conductance were recorded continuously during the expectation (pre-stimulus) and post-stimulus periods. The first was a single trial of a presentiment condition in which the upcoming stimulus was either an image of a pair of shoes (from the IAPS database - a non-arousing image) and no sound, or an erotic image (from the IAPS database - an arousing image) and the sound of a couple having sex (from the IANDS database - an arousing sound). The second was a single trial of a guessing condition in which the participant was asked to guess which of two pictures of birds would appear after the participant made his/her guess. The objective was to determine which of these two single-trial paradigms produced the largest physiological presentiment effect. We used 48 women and 46 men, and after removing outliers

(participants with mean SC or HR values more extreme than the mean +/- 3 x SD), we found a significant heart rate effect and a borderline skin conductance effect for women in the first (erotic image) condition, and a trend towards a significant skin conductance effect for men in this condition. In the second (guessing) condition, we found no significant effects for men or women.

Based on these results, we decided that the erotic image condition was either the most successful at generating presentiment effects, or that its position as the first experiment performed by the participants made it the most effective. Either way, we felt it would be best to do the formal single-trial experiment using only the erotic image paradigm. This formal experiment is just now underway, and it is this formal experiment that we are registering.

However, as we have decided to use smart phones to record data and we cannot use pornographic images on smart phones, we have attempted to imitate the excitement and active nature of the erotic image task in the formal experiment without including erotic images. Thus we will tell participants that 50% of them will win a prize (extra payment), and if the participant is a winner, visual and auditory displays will inform them of this. If the participant is a loser, they will still get paid a small amount, but not receive the comparatively large additional payment.

4. A statement or list of the specific hypothesis or hypotheses being tested, and whether each hypothesis is confirmatory or exploratory.

The specific hypothesis for the formal experiment is that there will be a significant gender x stimulus type interaction for the mean heart rate (HR) measure. This hypothesis is based on our previous result from the motivating analysis of our guessing study, as well as our analysis of the pilot data for the erotic image condition -- in both cases, male and female presentiment responses for heart rate measures were in opposing directions. This hypothesis is exploratory.

We have another exploratory hypothesis that HR and gender together can predict the upcoming stimulus type at a rate above chance performance.

5. The planned number of participants and the number of trials per participant.

Each participant performs 1 trial of the condition. We will run as many men and women in the experiment as it takes until we have clean data from 150 men and 150 women (N=300). Clean data are defined as data for which there is a continuous heart beat recording during the 30 seconds of the experiment, missing no more than 3 beats, and there is clean recording of the times of the initial button press, the offering of the feedback, and the end of the experiment. This definition of clean data arises because getting heartbeat timing data is relatively difficult in the Android development environment, with many possible hardware types. Thus, based on our experience using a commercial version of the experimental app, we need to clarify that the end goal is to get clean data from 150 men and 150 women. We will keep **all data** so they can be examined by a third party if desired.

6. A statement that the registration is submitted prior to testing the first participant, or indicating the number of participants tested when the registration (or revision to the registration) was submitted.

No participants have yet been run in this experiment.

In addition to the minimum content above, further information is highly recommended—particularly for well-planned confirmatory experiments. The additional information includes some or all of the following:

7. The specific statistical test method that is planned for each hypothesis, including which statistical test will be used, whether the unit of analysis is the participant or the individual random event, what *p* value (or confidence interval level) is significant, whether the statistical test (or confidence interval) is one or two-tailed, and any adjustment for multiple analyses. For example, "to analyze overall psi, a *z*-score binomial test with continuity correction will evaluate whether the overall rate of direct hits for all trials in the experiment is greater than 25%, with significance set at $p \leq .05$ one-tailed," or "the difference between the two conditions will be analyzed with a two-sample *t*-test with the number of hits for each participant as the unit of analysis and significance set at $p \leq .05$ two-tailed." (This information can be included in section 4 above.)

To test our hypotheses, we will use baselined heart rate data (see item 10 for baselining information) averaged over the 10-s period preceding stimulus onset. Based on the whole data set (N=300), participants with HR means beyond 3 x standard deviation of the data set +/- the mean of the data set will be removed as outliers. Once outliers are removed, we will perform the following statistical analyses on the remaining mean values (all tests will be two-tailed planned comparisons; alpha = 0.05 for each):

* A 2 x 2 (gender x stimulus type) ANOVA performed on HR means to test the exploratory hypothesis of a main effect of stimulus type.

* Two separate t-tests -- one on HR data from men, one on HR data from women to test the exploratory hypotheses that for both men and women, HR pre-stimulus responses will differ between stimulus types.

* Pattern classification analysis using second-by-second means of 18 seconds of pre-stimulus HR data for each participant (baselined to -20 to -18 seconds), 10 seconds of pre-stimulus HR data for each participant (baselined to -12 to -10 seconds), in addition to 45 bins of FFT data (obtained from the 18 seconds of pre-stimulus HR recording), plus participant gender and age, to create 75 factors to be used as independent variables in a pattern classifier that attempts to predict the upcoming stimulus type. This analysis will test the exploratory hypothesis that HR, FFT of HR, and gender can predict upcoming stimulus type above chance performance by comparing distributions of classification generalization error rate across the original data set and a scrambled version of the same data set using signal-detection statistics.

8. The power analysis or other justification for the number of participants and trials.

A power analysis on the pilot data determined that the effect was small enough that assuming about half of the participants win and about half lose, we will need about 120 participants of each gender to be 80% sure of seeing the effect. We use slightly more (150 of each gender) to ensure power and to account for outliers.

9. The methods for randomization in the experiment.

We use the pseudo-random number generator that is built in to the iPhone. Because there is only one trial, we feel that the RNG is not a concern (there is no order effect to be concerned about because there is only one trial, and participants are not run on a regular schedule so there is no particular time of day effect vis-a-vis the pseudo-RNG seed). The stimulus will be selected at random, just before it is presented.

10. A detailed description of the experimental procedure.

Before the experiment, participants are told that we are examining how physiology changes with expectation of an imminent event. In the consent form, we also tell them that there's a 50% chance that after pressing a button and seeing a visual 20-second countdown, they will see an image and hear a sound that they have won a prize (extra payment), and a 50% chance that they will see an image telling them that they have lost. After the participant agrees to the consent form and records his/her gender, age, and mTurk ID, a practice period begins in which the participant learns to position his or her finger over the webcam so the phone can pick up the heart rhythm. After successful practice, the app asks them to pick which of four colors will be the “winning color.” After this is done, a 20-second countdown begins. The experimenter does not know which stimuli will be picked (nor will the experimenter know until analyzing the data). a proprietary algorithm created by Mossbridge Institute, LLC for a different purpose (a commercial app) will be used to obtain HR data via the phone’s web cam.

After physiology data collection is started, the countdown begins. Once this countdown is over, the software uses the built-in pseudo-RNG to determine which of the two visual stimuli to present. If the choice is “win”, a happy sound is presented immediately, at the same time as a colorful image. If the choice is “lose”, there is no sound, there is a drab grey background, and the participant is informed that they lost. Stimuli are on the screen for 3 seconds. After the stimuli are displayed, participants are asked to remain for at least 10 more seconds so they may be paid. Data are recorded and sent to a database, and are coded according to a token entered into the database (the mTurk ID of the participant).

As we have done previously (Mossbridge et al. PA 2011), after using an automatic artifact-rejection algorithm, heart rate data are baselined to the average of the values –12 to -10 seconds before stimulus onset.

Note that for the classifier portion of the analysis, a copy of the HR data will be baselined to -30 to -28 seconds before stimulus onset.