# Study Registration for the
# Koestler Parapsychology Unit Study Registry

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

Experimenter Effect and Replication in Psi Research

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

Marilyn Schlitz, PhD, Senior Fellow, President Emeritus, Institute of Noetic Sciences
marilyn@noetic.org.
Arnaud Delorme, PhD, Visiting Scholar, Institute of Noetic Sciences
adelorme@noetic.org
Daryl Bem, PhD, Professor Emeritus, Cornell University
d.bem@cornell.edu

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

The proposed study seeks to study the replication problem in parapsychology through the examination of experimenter belief in psi. The meta-study involves an international collaboration of teachers, student experimenters, and experimental volunteers, who will make use of a standardized psi protocol developed by Daryl Bem that has been the focus of several recent replication attempts and that allows for a systematic collection of data under well-controlled conditions. In particular, Bem's studies were designed to be simple and transparent, requiring no instrumentation beyond a desktop computer, taking less than thirty minutes per session, and requiring statistical analyses no more complex than a t-test across sessions or participants. Specifically, the replication protocol will test the retroactive priming aspect of experiment 4 of Bem, 2011.

Each experimental session consists of 40 trials. In each trial an image is randomly selected and displayed, followed by a randomly selected incongruent or congruent priming word. Experimental volunteers will be instructed to identify images as "pleasant" or "unpleasant" as quickly as they can; after participants respond the priming word flashes briefly (at which point they do nothing until the next image shows). A total of 20 "unpleasant" and 20 "pleasant" images followed by a priming word (20 congruent and 20 incongruent) will be shown. To also test for experimenter beliefs, our study will first question experimenters about their belief in or receptivity to psi, based on an opportunistic selection of experimenters. The

responses to these questions will be the independent variable. Experimenters will receive a prerecorded, web-based experimenter orientation and training (see http://www.youtube.com/watch?v=cl7Jag_xGJY). Experimental volunteers will be recruited and tested on the retrocausal priming experiment by the experimenter and will also answer a short battery of questions about their beliefs about psi phenomena prior to testing. The results of the psi task – an analysis of the "pleasant" vs. "unpleasant" selection response time – will be the dependent measure for both the psi replication attempts and for the experimenter beliefs.

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

The proposed experiment will first be a large scale replication of one of D. Bem's experiments from his landmark article "Feeling the future". We will also test the hypothesis that the belief of the experimenter will have an impact on the outcome of a psi task. In particular, it will explore one of the most controversial aspects of psi research: precognition. The psi experiment tests the confirmatory hypothesis that memory can "work both ways" by testing whether words can influence the perception of an image —even if the image presentation takes place before the word is given.

This experiment will investigate one confirmatory and two exploratory hypotheses. (1-confirmatory) Replicating the previous study by Bem, response time will be shorter for trials with congruent words than for trials with incongruent words. (2-exploratory) The response time effects (differences) will be greater for experimenters with positive expectations (belief in psi effects) about the experimental outcome than for experimenters with negative expectations (lack of belief in psi effects). (3-exploratory) The response time effects will be greater for participants with positive beliefs/expectations about psi than for participants with negative beliefs/expectations about psi.

The hypotheses regarding experimenter and participant belief/expectancy is a planned exploratory analysis.

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

The protocol will involve three levels: (1) Teachers (T) who will recruit student experimenters, (2) Student Experimenters (SE) who will receive a standardized training in the experimental procedure, (3) participants (P) in the study who participate in the psi task.

32 SE will be identified and recruited to participate in the study. Each SE will recruit 16 participants (P) for a study about precognition. This will be a total of 512 subjects who will each perform the psi task once. Each participant will respond to the 5 questions (Annex 1) and be presented with 40 images followed by 40 congruent and incongruent words.

*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.*

> We confirm that this registration is submitted prior to the testing of the first participant. The total number of participants will be 512.

> To prevent criticism against potential falsification of data on our part, all student experimenters will be asked to send the data by email simultaneously to us as well as to a third neutral party not involved in the experiment (Caroline Watt at caroline.watt@ed.ac.uk).

*7. The specific statistical test method that is planned for each hypothesis, including any adjustment for multiple analyses.*

> Because response-time data are positively skewed, each response time (RT) will be usually transformed prior to analysis using either an inverse transformation (1/RT) or a log transformation (log RT). Trials yielding very short or very long response times will be considered to be spurious outliers and will be excluded from the analysis. Ratcliff (1993) suggested using more than one cutoff criterion to ensure "that an effect is significant over some range of nonextreme cutoffs" (p. 519). We will run four analyses, using both data transformations and two different cutoff criteria for long response times, 1,500 ms and 2,500 ms. In all cases we will use parametric statistics to assess significance.

> To provide an analysis that avoids distribution assumptions, we will also compute the percentage of participants who had positive priming scores in each condition, evaluated by an exact binomial test.

> To test if the effect observed by Bem (2011) is reproducible, we will simply be pooling all participants run by all student experimenters. We will consider all the data collected as if originating from one single large experiment.

> To test the experimenter effect, we will assess if SE as well as participants' belief in psi phenomena (as assessed with the short questions in Annex 1) influence the results. We will be using parametric statistics using the psi belief of the participant and the experimenter as independent variables and effect size as dependent variable (3x3 ANOVA design). We expect to see both an effect of the participant belief and an effect of the experimenter belief on the results.

> Specifically, the analyses for the three hypotheses identified in item 3 are as follows:

Hypothesis 1: Following Bem, 2011, the analysis for overall psi will be:
1. One-tailed tests with significance set at $p=.05$;

2. The unit of analysis is the participant, with the measure for each participant being the mean response time for incongruent trials minus the mean response time for congruent trials;
3. Four statistical tests will be done using two response time data transformation (1/RT and log(RT)) combined with two outlier cutoff criteria (exclude trials with response times >1500 ms and > 2500 ms);
4. Data transformations will be applied to the raw response time for each trial.
5. Trials with errors in judging the image as pleasant or unpleasant will be excluded;
6. Trials with response times <250 ms will be excluded;
7. Participants with judging errors on 25% or more of the trials will be excluded;
8. Each of the four analyses will use a single-mean t-test combining all non-excluded participants and testing the one-tailed alternative hypothesis that the measure of response time difference is greater than zero;
9. To avoid distributional assumptions, overall psi will also be evaluated with an exact binomial test evaluating the one-tailed alternative hypothesis that the percent of participants with a positive measure of response time difference is greater than 50%.


Hypotheses 2 and 3: For the analyses for experimenter and participant expectancy/beliefs:
1. Four tests for each hypothesis, using the dependent variable and the exclusion criteria as described above will be used;
2. Five questions to assess belief in or receptivity to psi will be used for classifying experimenters and participants beliefs. The five questions are shown in Annex 1.
3. The five questions will be used for classifying experimenters and participants' beliefs. We will take the sum of the responses for questions 3 and 4 (the sum will be from 0 to 8) to assign participants on one hand and experimenters on the other hand to 3 groups:
   a. The positive group will be those who score in the upper 33% on the 0 to 8 scale (33$^{rd}$ percentile on the distribution of scores that are actually obtained in the experiment).
   b. The negative group will be those who score in the lower 33% on the 0 to 8 scale.
   c. The middle group will be everyone else.
4. The analysis will use a 3x3 ANOVA (experimenter beliefs as one factor and participant beliefs as the other factor);
5. Significance will be based on the main effect for each factor with significance set at $p=.05$.


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

Bem set 100 as the minimum number of participants/sessions for each of the experiments reported in this article because most effect sizes reported in the psi literature range between 0.2 and 0.3. If d = 0.25 and N = 100, the power to detect an effect significant at .05 by a one-tailed, one-sample t test is .80. Cohen's classic

1988 book (Cohen, 1988) on power analysis has tables of effect sizes for various experimental and statistical designs.

The two priming studies in Bem's 2011 article (Experiments #3 and 4) actually achieved effect sizes of 0.26 and 0.23, respectively. Both experiments had 100 subjects. The current study uses a total of 512 in an attempt to increase the power of the main statistical analysis.

*9. The methods for randomization in the experiment.*

There is no randomization procedure when selecting subjects – subjects are not assigned to different groups.

Computer randomly selects priming words to display after the image. This random selection is performed using the random number generator associated with the presentation software. The presentation software is Real Basic. In this experiment, randomizing was implemented by Marsaglia's PRNG algorithm.

*10. A detailed description of the experimental procedure.*

The protocol will involve three levels: (1) Teachers (T) who will recruit student experimenters, (2) Student Experimenters (SE) who will receive a standardized training in the experimental procedure, (3) participants (P) in the study who participate in the psi task.

Drawing on the existing professional networks of teachers of the investigators, 32 SE of research classes will be identified and recruited to participate in the study (we are confident we can recruit at least 32 Student Experimenters). They will be selected based on their interest in the study, not by their pre-existing belief in a particular outcome for the replication attempt. Each SE will recruit 16 participants (P) for a study about precognition. This will be a total of 512 subjects. The psi experiment tests the hypothesis that memory can "work both ways" by testing whether words can influence the perception of an image —even if the image presentation takes place before the word is given.

Each person at the three levels of the study (teachers, student experimenters, participants) will be assessed for their baseline belief in psi phenomena and their belief in psi phenomena the replication study; we will make use of 5 simple questions (see Annex 1) to assess belief in psi.

SE's will be given a standardized training in the experimental protocol prior to initiating their experiment. A web-based short video will instruct them about the experiment protocol for data collection. SE will download a Windows or Mac based experimental program that they will install on a computer to collect data. The program will be identical for all SE. A short survey will be used to assure that the SE's understand the procedure.

Based on a pre-existing study design and replicable experimental series reported by Bem (2011), participants will first be shown images (retroactive priming part of experiment 4 of Bem, 2011). Participants will be told that a word will be flashed on the screen just after they made their judgment of the picture.

Upon entering the laboratory, the participant will be told, "This experiment tests for ESP by administering several tasks involving common everyday words. The experiment is run completely by computer and takes less than 20 minutes. The program will give you specific instructions as you go. At the end of the session, I will explain to you how this procedure tests for ESP."

Participants will be seated in front of the computer. After they have responded to the 5 questions shown in Annex 1 and gone through a 3-min relaxation procedure (see Bem, 2011), they will start the retroactive priming task.

20 positive and 20 negative pictures will be drawn from the IAPS set. 20 positive and 20 negative prime words will be used and a prime will be randomly selected on each trial after the participant has responded to the picture. As a result, congruent trials and incongruent trials are randomly sequenced and do not necessarily occur in equal numbers. This makes it virtually impossible for participants to anticipate the type of trial coming up by knowing the types of trials that have already occurred.

Following completion of the experimental session, the SE will transfer the data files by email to the Institute of Noetic Sciences as well as to a third neutral party. Data files will be archived. When all SE have contributed their data files, they will be analyzed.

**References:**

Bem, D. (2011) Feeling the Future: Experimental Evidence for  Anomalous Retroactive Influences on Cognition and Affect. *Journal of Personality and Social Psychology, 100,* 407-425.

Cohen, J (2005) Statistical Power Analysis for the Behavioral Sciences, Psychology press.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. Psychological Bulletin, 114, 510–532.

**Annex 1: questions about belief in psi**
(numbers next to each response will not be shown to subjects)

1) I often enjoy seeing movies I've seen before
> Very untrue
> Untrue
> Between true and untrue
> True
> Very true

2) I get bored easily
> Very untrue
> Untrue
> Between true and untrue
> True
> Very true

Extrasensory Perception (ESP is the Reception of Information Without the Use of the Known Senses. It includes:
- Telepathy: ESP of the Thoughts of Another Person.
- Clairvoyance: ESP of Hidden Objects of Distant Events. (Also called Remote Viewing.)
- Precognition: ESP of Future Events.

3) Do you believe ESP exists?
> Definitely does not   (0)
> Probably does not   (1)
> Don't know   (2)
> Probably does   (3)
> Definitely does   (4)

4) Have you had any experiences that you believe were genuine ESP?
> Definitely not   (0)
> Probably not   (1)
> Maybe   (2)
> Probably yes   (3)
> Definitely yes   (4)

About yourself
5) Have you ever practiced any form of meditation, self-hypnosis, relaxation exercises, or biofeedback?
> No never
> Only a few times
> Occasionally
> Regularly in the past
> Regularly now