HW 1 Write Up

In a test by Bem, designed to measure the existence of premonition, participants’ ability to predict future events was compared to the rate of chance at which those events were presented. Participants seated at a computer were instructed to view trials of randomly intermixed photos of erotic or non-erotic content. Bem found that participants were successfully able to identify the position of the erotic pictures significantly more often than chance. In an attempt to replicate Bem’s results, these data have been collected independently for analysis. The first test, following his sample size (N=100) and one-sided t-test design, yielded a non-significant result (p=0.9979, t(99)=-2.93). The second test used a sample size calculated from a power analysis based on Bem’s mean effect size (N=271) and groups were compared using a two-sided t-test, resulting in non-significance (p=0.8789, t(270)=0.15). Both tests maintained Bem’s mean effect size of d=0.22.

If our theory suggested that erotic events shocked people, disrupting their psi ability, the directionality of our one-tailed t-test would have been reversed, resulting in a significant finding (p=0.002, t(99)=-2.3). However, our results for the second replication would remain the same, since the two-tailed test already includes changes in abilities occurring on either end of the distribution.

If a two-tailed t-test had been conducted on the first replication, a significant result would have been found (p=0.004, t(99)=-2.93). This is because the initial one-tailed t-test failed to capture the effect on the negative side of the distribution. Two-tailed tests do raise the critical value to reach significance, but it allows you to capture results on either side.

Power can be increased by increasing the number of subjects, lowering the variance, and/or increasing the effect size. With an under-powered replication there is more variance associated with the smaller N, causing a Type 1 error. The first replication had a larger d than the high powered replication because the N was increased from 100 to 271, which along with increasing the sample size, helped to reduce variance in the sample.

In Bem’s design, he used a one-sample t-test to compare the average response to chance. If Bem had instead used a two-sample test, he would have been comparing the two group means. If the groups had done similarly well, he would have failed to reject the null hypothesis. If the one group had correctly predicted the future event significantly more than the other group, he would have been able to reject the null hypothesis.