Analysis Plan for Risen & Gilovich RRR

March 17, 2017

Contact: Maya Mathur (mmathur@stanford.edu)

Notation

- Y: in the classroom scenario, subject's perceived likelihood of being called on
- X: tempting fate manipulation (0/1)
- L: cognitive load manipulation (0/1)
- S: replication site similarity manipulation (0=MTurk / 1=similar university; 3 nominal levels in Models 1'-3' as described below)
- A: achieved cognitive load (continuous; sum of secondary questions regarding effort and difficulty of the cognitive load task)
- P: academic pressure (continuous; response to secondary question regarding importance of answering questions correctly in class

1 Primary Analysis Models

Each primary analysis model will be fit only among United States sites designated *a priori* as being similar academic environments to Cornell University, the site of the original study. We defined "similar sites" as U.S. colleges/universities with median SAT scores of >90th percentile (with percentiles based on individual student scores nationally) on all sections of the test (or in total if section scores are not available). Cornell is in the 95th percentile. We also recruited from universities that did not meet these criteria, called "dissimilar sites".

Throughout, all models are linear mixed-effects models with normally distributed, independent error terms fit via REML or ML using `R`'s `lme4` package. In each model, random effects reflect the maximal potentially identifiable model structure given the hierarchical study design. If model estimation fails to converge, we will refit the model by iteratively dropping individual random effects to yield the maximal model whose estimation converges. We will report (at minimum) point estimates, 95% confidence intervals, and p-values for all model parameters of interest. Inference will not employ multiplicity corrections for primary or secondary analysis models.

Primary analyses address two central research questions:

(Q1): Is the target effect of interest stronger in similar sites vs. MTurk?

(Q2): What is the estimated magnitude of the target effect of interest in similar sites? In the MTurk?

The target effect of interest is the $X * L$ interaction, per the selection criteria from RP:P. However, we also assess the main effect of X, which was also a central finding of Risen & Gilovich (2008).

1.1 Model 1

$$
\begin{aligned}
Y_i = \beta_0 + \beta_X X_i + \beta_L L_i + \beta_S S_i + & \quad \text{(fixed main effects)} \\
\beta_{XL} X_i L_i + \beta_{XE} X_i S_i + \beta_{LS} L_i S_i & \quad \text{(fixed 2-way interactions)} \\
\beta_{XLS} X_i L_i S_i + & \quad \text{(fixed 3-way interaction)} \\
\gamma_{0,i} + \gamma_{X,i} X_i + \gamma_{L,i} L_i + \gamma_{XL,i} X_i L_i & \quad \text{(random site intercepts and slopes)} \\
+ \epsilon_i &
\end{aligned}
$$

Reporting and statistical inference will focus on the following parameters and linear combinations:

- β_{XLS}: effect of site similarity on the $X * L$ interaction that is the target effect of interest **(Q1)**
- β_{XS}: effect of site similarity on the main effect of tempting fate **(Q1)**
- β_{XL}: the magnitude of the interaction effect of interest within MTurk **(Q2)**
- $\beta_{XLS} + \beta_{XL}$: the magnitude of the interaction effect of interest among the similar sites **(Q2)**
- β_X: the magnitude of the main effect of interest in MTurk **(Q2)**
- $\beta_{XS} + \beta_X$: the magnitude of the main effect of interest among the similar sites **(Q2)**

Additionally, as a sensitivity analysis, we will refit the exact ANOVA model used in Risen & Gilovich (2008) among only the similar sites. (The motivation for fitting our more complicated model is its ability to estimate to address the ML5 question regarding protocol similarity.)

2 Secondary Analysis Models

2.1 Models 1'-3'

These models will include all sites. We will therefore redefine S as a nominal (non-ordinal) variable with three categories: similar, dissimilar, and MTurk (defined in the Revised Protocol).
We will refit primary Models 1-3, now among all sites. Reporting will center on the same parameters as previously (with additional terms reflecting the third level of the similarity variable).

2.2 Model 4

If Models 1'-3' suggest that results differed for similar and dissimilar sites, we will investigate a proposed mechanism via two additional moderated mediation models. Models 4 and 5 will again use only similar sites, as in primary analyses and will be estimated via Imai et al.'s `R` package, `mediation`.

Model 4 will be fit only among subjects with $L = 1$ (those assigned to cognitive load). It will assess the secondary, mechanistic hypothesis:

(Q3) Among subjects with $L = 1$, does achieved cognitive load (A) mediate the effect of S on Y, moderated by X, such that the indirect path through A is stronger when $X = 1$?

That is, for subjects who tempted fate and are under cognitive load, the $S \rightarrow Y$ pathway through achieved cognitive load is strong. For subjects who did not tempt fate and are under cognitive load, the $S \rightarrow Y$ pathway through achieved cognitive load is weak. (This is because, for subjects who are supposed to be under cognitive load, tempting fate increases the effect of achieved cognitive load on Y. Framed equivalently but in more intuitive terms, achieved cognitive load increases the effect of tempting fate on Y.)

```
med = lm(A ~ S*X, data=d[d$L==1,])
out = lm(Y ~ S*X + A, data=d[d$L==1,])
```

2.3 Model 5

Similarly, this model will assess the secondary, mechanistic hypothesis regarding the main effect of X:

(Q4) Academic pressure (P) mediates the effect of S on Y, moderated by X, such that the indirect path through A is stronger when $X = 1$.

That is, the academic pressure pathway (from $S \rightarrow Y$) is stronger for subjects who tempted fate vs. those who didn't. This is because the effect of academic pressure depends on whether one has tempted fate (or, reframed: the effect of tempting fate depends on academic pressure).

```
med = lm(P ~ S*X)
out = lm(Y ~ S*X + P)
```

3 Statistical Power

Analyses will be powered at least 80% to detect a 3-way interaction (β_{XLS}) representing a change in perceived likelihood of 0.75 standard deviations in primary analyses (using only similar sites). We will therefore attempt to collect at least $n = 320$ in similar sites and $n = 800$ on Mechanical Turk, but may increase the latter sample size if needed to compensate for under-enrollment in the similar sites.

The plot below shows simulated power using estimates from the RPP data as a function of Mechanical Turk sample size and total in-lab sample size across all similar sites (code).