

Supplementary Analyses for: “Registered Multisite Replication of Tempting-Fate Effects in Risen & Gilovich (2008)”

Contents

Descriptive Statistics and Plots	2
Cohen’s d and t-tests within each site	2
Interaction plots by site type	2
Cell means and standard deviations by site type	5
Sensitivity Analyses for Reported Results	5
Fit subset model counterpart to primary analysis model	5
Fit meta-analytic counterparts to primary analysis model	5
Combine all universities into one category	6
Refit original study’s ANOVA model	6

Descriptive Statistics and Plots

Cohen's d and t-tests within each site

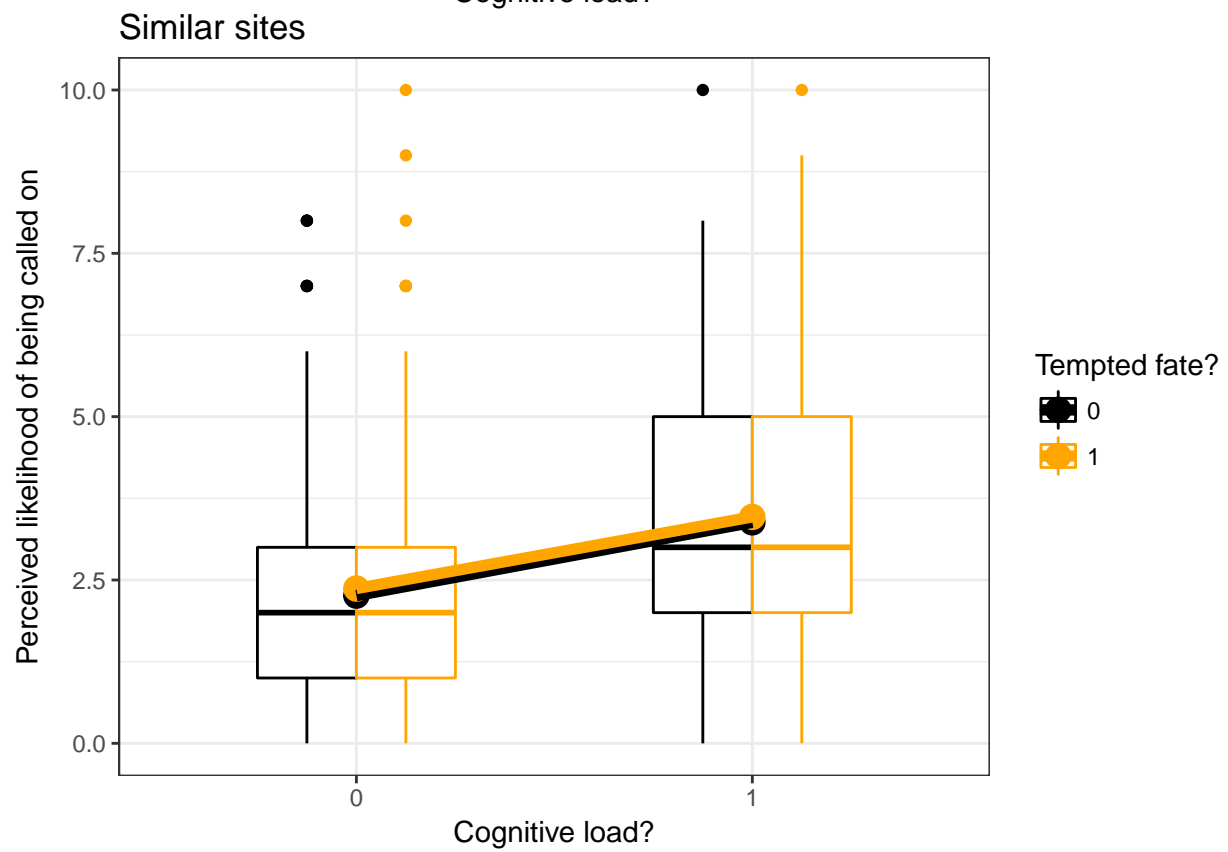
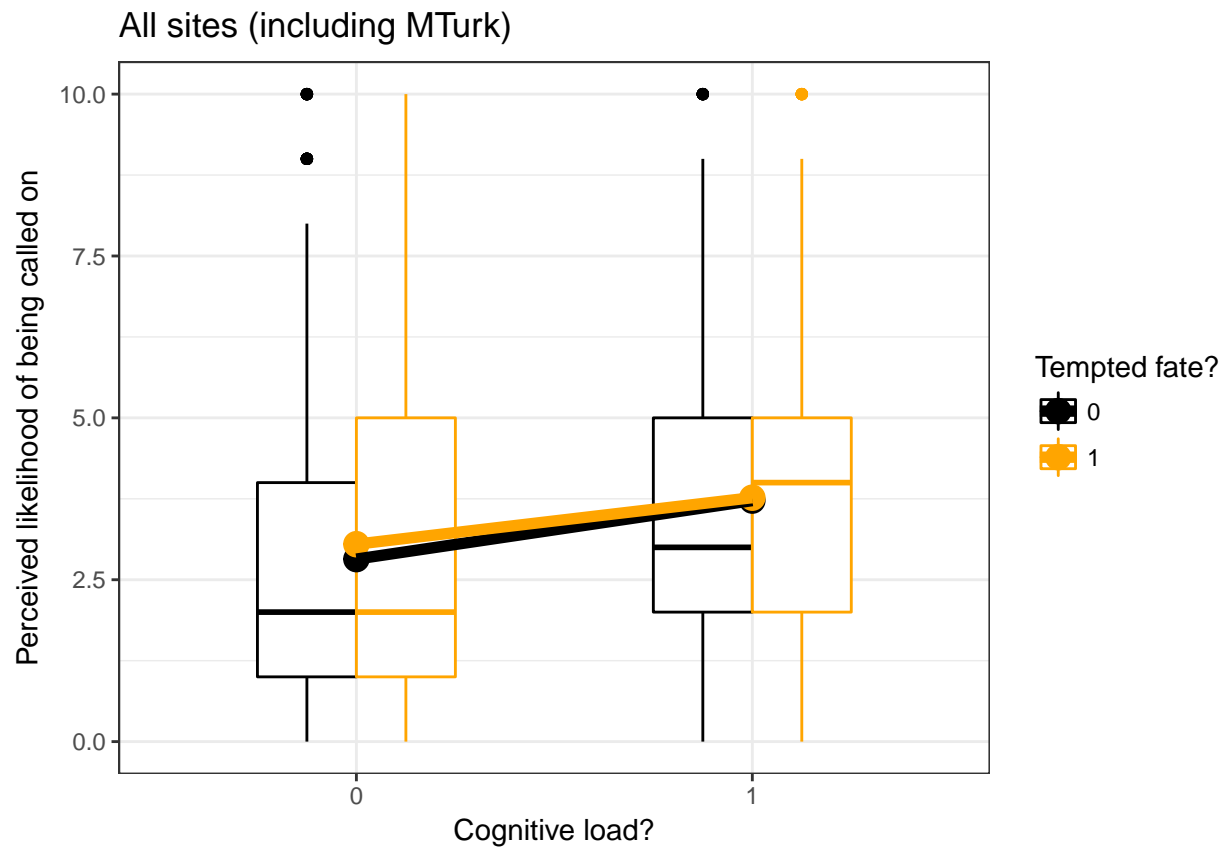
Per the preregistration, here we conduct additional within-site analyses that reproduce the original study's stratified analyses and effect sizes of tempting fate.

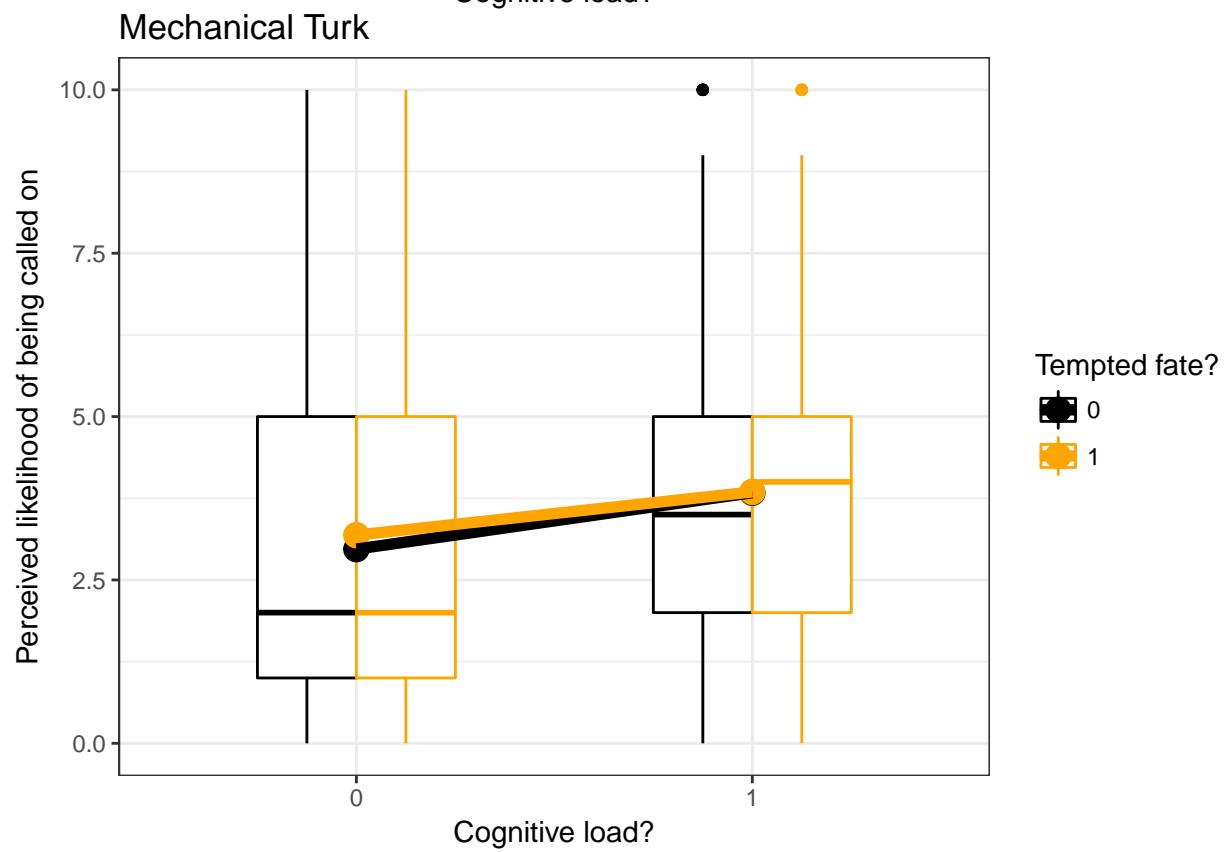
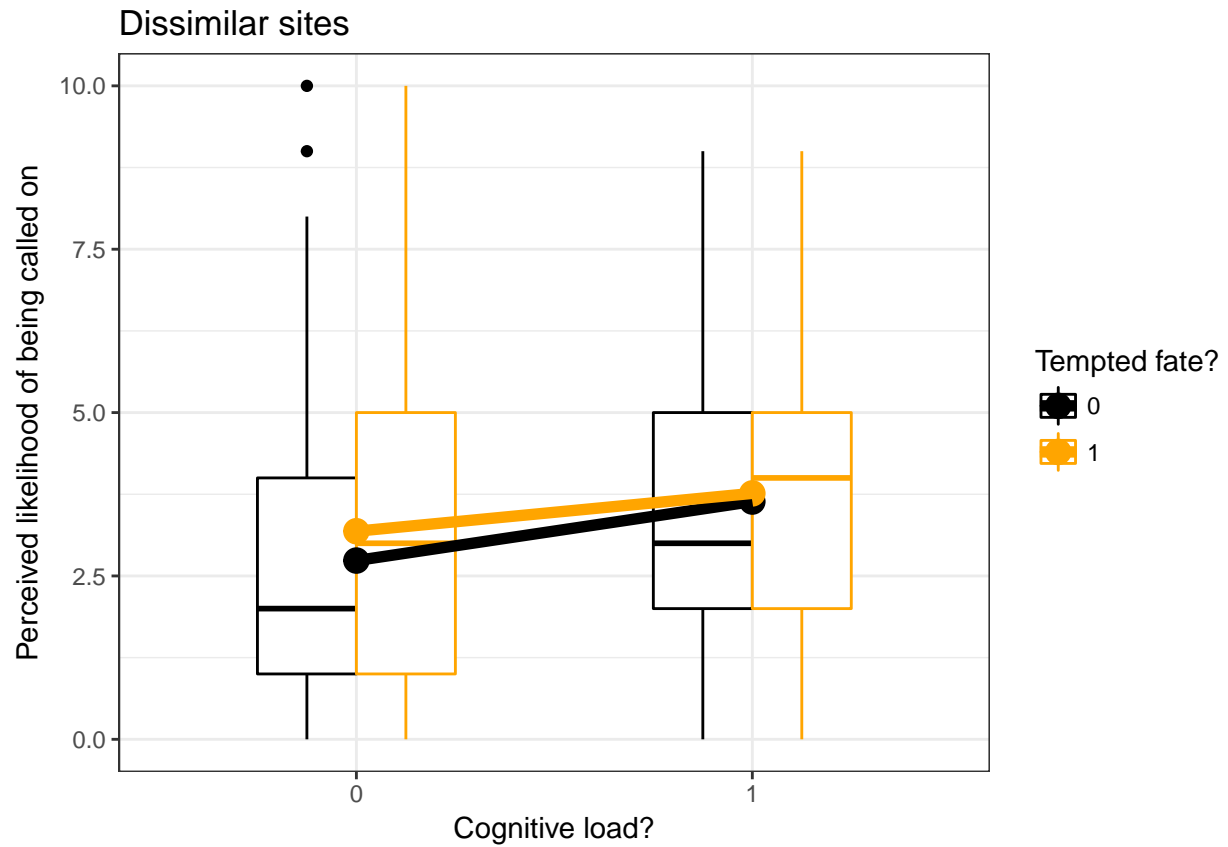
Table 1: Cohen's d and p-values for t-tests of the effect of tempting fate on perceived likelihood, stratified by cognitive load within each site.

Site	Cohen's d (no load)	p-value (no load)	Cohen's d (load)	p-value (load)
MTurk	0.09	0.07	0.01	0.92
U Penn	0.13	0.40	0.20	0.20
UCB	-0.11	0.55	-0.30	0.16
UVA	0.02	0.94	0.06	0.79
Stanford	0.40	0.22	0.20	0.59
Eotvos	0.50	0.00	0.03	0.84
KUL	0.06	0.81	0.58	0.03
UP	0.00	1.00	0.02	0.95
BYUI	0.04	0.90	-0.10	0.76
URI	0.03	0.91	0.02	0.95
RHIT	0.10	0.79	-0.02	0.97

Interaction plots by site type

Boxplots: medians and IQRs; lines: simple means by subset. These aggregated means and SDs pool across all sites within a group (similar, dissimilar, MTurk) and do not account for clustering by site.





Cell means and standard deviations by site type

Table 2: Means and SDs of perceived likelihood across all subjects within each site type (naively pooling all sites)

tempt	load	group	mean	SD
0	0	a.mturk	2.97	2.39
1	0	a.mturk	3.19	2.41
0	1	a.mturk	3.84	2.31
1	1	a.mturk	3.85	2.32
0	0	b.similar	2.26	1.89
1	0	b.similar	2.37	1.96
0	1	b.similar	3.38	2.16
1	1	b.similar	3.47	2.19
0	0	c.dissimilar	2.74	2.02
1	0	c.dissimilar	3.18	2.19
0	1	c.dissimilar	3.64	2.10
1	1	c.dissimilar	3.76	2.01

Sensitivity Analyses for Reported Results

Fit subset model counterpart to primary analysis model

Instead of fitting a model that includes both MTurk and similar sites with an interaction of site type, we fit a model to only the subset of similar sites.

```
m1.temp = lmer( lkl ~ tempt * load + (tempt * load | site), data = b[ b$group == "b.similar", ] )
CI.temp = confint( m1.temp, method = "Wald" )
```

In the primary model, the estimated main effect was 0.11 with 95% CI: (-0.34, 0.56), whereas in the present subset model, it is 0.13 with 95% CI: (-0.33, 0.60). Also, in the primary model, the estimated interaction effect was -0.02 with 95% CI: (-0.68, 0.63), whereas in the present subset model, it is -0.03 with 95% CI: (-0.66, 0.6). These results are similar.

Fit meta-analytic counterparts to primary analysis model

Instead of fitting a mixed model to observation-level data, we fit a random-effects meta-analysis to the point estimates using the Paule & Mandel heterogeneity estimator and the Knapp-Hartung standard error adjustment. For the main effect:

```
meta.main = rma.uni( yi = site.main.est, vi = site.main.SE^2,
                    data=sites[ sites$group == "b.similar", ],
                    measure="MD", method="PM", knha = TRUE )

p_orig( orig.y = yi.orig.main, orig.vy = vyi.orig.main,
        yr = meta.main$b, t2 = meta.main$tau2,
        vyr = meta.main$vb )
```

```
##
```

```
## If the original study were statistically consistent with the replications, the probability of an est.
```

```
## [1] 0.06872944
```

In the above mixed model, the estimated main effect and heterogeneity in similar sites was $\widehat{M} = 0.13$ and $\widehat{V} = 0.06$ compared to $\widehat{M} = 0.13$ and $\widehat{V} = 0$ in the meta-analysis. They agree very closely. P_{orig} is a bit lower due to the lower estimated heterogeneity here.

For the target interaction effect:

```
meta.int = rma.uni( yi = site.int.est, vi = site.int.SE^2,
                    data=sites[ sites$group == "b.similar", ],
                    measure="MD", method="PM", knha = TRUE )

p_orig( orig.y = yi.orig.int, orig.vy = vy.orig.int,
         yr = meta.int$b, t2 = meta.int$tau2,
         vyr = meta.int$vb )
```

```
##
## If the original study were statistically consistent with the replications, the probability of an est.
## [1] 0.04174708
```

In the above mixed model, the estimated interaction effect and heterogeneity in similar sites was $\widehat{M} = -0.03$ and $\widehat{V} = 0.06$ in the mixed model compared to $\widehat{M} = -0.02$ and $\widehat{V} = 0$ in the meta-analysis. They agree reasonably closely.

Combine all universities into one category

In the planned secondary analysis model including all universities, similar and dissimilar sites were treated as separate categories. Here, they are combined into one category.

Name	Estimate	CI	pval
Tempt main effect within MTurk	0.21	[-0.28, 0.71]	0.40
Tempt main effect within university sites	0.28	[-0.08, 0.63]	0.12
Effect of university site vs. MTurk on tempt main effect	0.06	[-0.55, 0.67]	0.84
Tempt-load interaction within MTurk	-0.20	[-1.01, 0.60]	0.62
Tempt-load interaction within university sites	-0.17	[-0.69, 0.35]	0.52
Effect of university site vs. MTurk on tempt-load interaction	0.03	[-0.93, 0.99]	0.95

Refit original study's ANOVA model

The original study used two-way ANOVA to test for the main effect and interaction. Per our preregistered protocol, we also reproduce this model as a secondary analysis here. Since this model is statistically equivalent to the regression models presented in the main text, this is simply a different way of presenting the contrasts. The results are qualitatively similar to those in the main text.

```
# with standard ANOVA mean contrasts and sequential decomposition
# main effect: half the effect of tempting fate vs. not tempting fate when not under load
summary( aov( lkl ~ tempt * load, data = b[ b$group == "b.similar", ] ) )
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## tempt      1    1.4    1.36    0.325    0.569
## load      1  230.5  230.53  55.009 3.25e-13 ***
## tempt:load  1    0.0    0.03    0.007    0.933
## Residuals 750 3143.0    4.19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

# with contrasts vs. 0 and marginal SS decomposition
# main effect: effect of tempting fate when not under load
summary( lm( lkl ~ tempt * load, data = b[ b$group == "b.similar", ] ) )

```

```

##
## Call:
## lm(formula = lkl ~ tempt * load, data = b[b$group == "b.similar",
##    ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.4655 -1.3829 -0.3829  1.5345  7.6311
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.26131    0.14512  15.583  < 2e-16 ***
## tempt        0.10763    0.20347   0.529   0.597
## load         1.12155    0.21215   5.287 1.63e-07 ***
## tempt:load   -0.02497    0.29905  -0.083   0.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.047 on 750 degrees of freedom
## Multiple R-squared:  0.06872,    Adjusted R-squared:  0.06499
## F-statistic: 18.45 on 3 and 750 DF,  p-value: 1.485e-11

```