Why People are Reluctant to Tempt Fate: Replication of Study 6 and Extension

Valeria Orozco Castiblanco
Universitat Pompeu Fabra Barcelona
Master's in Cognitive Systems and Interactive Media
Barcelona, Spain
valeria.orozco01@estudiant.upf.edu

Andrei Ghiuru
Universitat Pompeu Fabra Barcelona
Master's in Cognitive Systems and Interactive Media
Barcelona, Spain
andrei.ghiuru@tuta.io

Nicole Ubillius
Universitat Pompeu Fabra Barcelona
Master's in Cognitive Systems and Interactive Media
Barcelona, Spain
nicole.ubillus01@estudiant.upf.edu

***Abstract*—The present research focused on the replication of the sixth experimental study (Study 6) from the paper "Why People are reluctant to tempt fate", published by authors Risen and Gilovich in 2008. The replication was developed with the aim of supporting previous results from the authors that had revealed the dual thinking process (System 1 and System 2) to be the cause of the belief that it is bad luck to "tempt fate". A higher perception in the likelihood of a negative outcome had been predicted when the participants were under cognitive load while performing the task, in other words, when System 2 was loaded. A close replication of the study was made, and cognitive load did not show to have an effect on the perceived likelihood of a negative outcome after tempting fate. Having failed to replicate the results of the original study, a further methodological extension was proposed. The extension of the study followed the same methodology proposed by the authors, with three modifications: (1) two new variables were added (Importance and Relevance) to control if whether the participants cared or not about the tempting fate scenario of the student, (2) changes in sampling, and (3) the method by which participants were loaded cognitively, which will be described in detail in the method section. The results showed that participants who had rated high in Importance (i.e. those who would have felt bad had they been called by the professor without having read the required reading), when loaded cognitively, had a greater perception of a negative outcome. This is regardless of the condition of tempting fate or not, which is similar to a replication done by Maya Matur in 2016. The other main effect from the original study (tempting fate has an effect on perception of negative outcomes) and the interaction between cognitive load and tempting fate could not be replicated.**

Keywords—replication, intuition, magical thinking, tempting fate, dual thinking.

I. INTRODUCTION

In the age of science and technology, it could be argued that magical thinking is a disappearing phenomenon. Surprisingly, it is not. Magical thinking has been described as a universal, common trait in people [1], one that was previously seen as abnormal or attributed to mental deficiencies. Due to the change in the conception of the term, a more suitable and less pathologizing definition of the concept has appeared. Magical thinking has been defined as the belief that certain actions can influence objects or events when there is no empirical causal connection between them [2]. Superstitious and magical beliefs are not just scientifically wrong, but scientifically impossible, according to Risen [3].

An attempt to explain the human psychological tendency to believe in superstitions has been done under the dual process of thinking provided by Kahneman [4], in which he describes two exclusive ways of thinking that people use interchangeably to give responses to everyday life situations: System 1, fast thinking, intuitive, automatic and reactive, and System 2, defined as slow thinking, conscious, deliberate, effortful and statistical.

According to Risen [3], under the dual process model described above, the emergence and maintenance of a magical belief is closely related to the level of engagement of System 2 in correcting the initial intuition given by System 1. For instance, believing it is more likely to rain when someone takes an umbrella with them, than when they don't, is a judgement with no empirical causal connection generated by System 1. It is then the role of System 2 to correct the initial judgement, by bringing to consciousness the fact that such connection is impossible. However, correcting these ideas does not happen equally for everyone, there are certain features that prompt correction of System 1's incorrect judgements described by the author [3]. Amongst them, the ability and motivation to be rational, which are related to cognitive ability and education on one hand, and the motivation and incentives to be rational on the other.

Personal differences can lead to some people being more motivated to think rationally than others, leading them to be less likely to engage in superstitious beliefs, just like those who are more educated. However, magical thinking is not limited to certain populations and the tendency towards believing in superstitious beliefs is not rooted in both motivation and education, rather for Risen [3] these variables add to the original layer, which is the dual process of thinking, and can help predict how likely it is that System 2 will engage or not in correcting an irrational assumption.

Some experimental approaches to why people tend to engage in magical beliefs even if they know they are irrational have been made under the dual process model. An example is Callan et al. [5] experiment, in which they explored how immanent justice, understood as the tendency to assume that misfortunes befalling others are the workings of justice, was present in people. For this, the authors created a between subjects design, and presented participants with two different situations. In both, participants were under some kind of cognitive load, either memorizing a two (light) or a nine digit (high) number before reading a situation about David, who had suffered an accident.

For one group, prior to the accident David had been having an adulterous affair with Susanne, a travel agent. For the other, he had been visiting the travel agency to purchase a vacation to Mexico with his family. Participants were asked to what extent did they believe David deserved to be in the accident, and to what extent they felt David's accident was the result of

his dealings with Susanne by rating on a Likert scale from 1 to 7, 1 being not at all and 7 being a great deal.

The results showed perceiving the accident as a deserved punishment was significantly stronger under high than low cognitive load. Additionally, the participants' perception of a predicted misfortune after David's conduct was significantly stronger under high than low cognitive load. Which suggests that, when under load, System 2 cannot adjust the incorrect assumptions that emerged from System 1, nor filter irrational judgements, implying that the dual system of thinking is indeed related to the judgement of a situation.

Another attempt to explain why people engage in magical thinking was made by Risen and Gilovich in 2008 [1]. Following a set of experiments, they explored the existence of the belief that, when tempting fate, a negative consequence or punishment would follow. The authors used McKean [6] definition of tempting fate, which is "to do something that is risky or dangerous".

After a series of experiments about the perceived likelihood of a negative outcome following a tempting fate situation, the authors arrived at the conclusion that those in their study did tend to be reluctant to tempt fate, and believed a negative consequence could emerge if they did it. Following the confirmation of a tendency to being reluctant to tempt fate, the authors engaged in explaining the underlying cause of the found tendency, by creating an new experiment.

From the set of experiments in the paper, Study 6 looked into the underlying cause of people being reluctant to tempt fate, the authors hypothesizing that since System 2 is the one responsible for correcting irrational assumptions made from System 1 way of thinking, when under load, the effect of perceived likelihood of a negative outcome after tempting fate would be significantly increased.

To assess the hypothesis, Risen and Gilovich designed an in-between subjects factorial design experiment, the participants, who were all undergraduate students from Cornell University, were assigned to either cognitive load condition, which meant counting while reading or, no cognitive load condition, which meant just reading. Afterwards, the participants were presented with a story where they were impersonating a student in a large class. The student had either tempted fate (hadn't read the required reading for the class) or not (had read the required reading). After no student volunteers to answer the professor's question, the professor decides to call on a student to answer the question randomly. The participant is asked how likely it is that they will be called on in a Likert scale from 0 to 10, assuming 0 as no possibility to be called and 10 as certainty of being called.

In the study, the authors were able to reject the null hypothesis and show a significant difference in terms of likelihood of a negative outcome when the participants were under cognitive load and tempting fate condition (M = 5.25, SD = 2.36), in contrast to the results when assigned to no load and no tempting fate conditions (M = 2.70, SD = 2.17). The results suggest that when System 2 is under load, people will have a lower tendency to correct irrational assumptions and therefore, it will be more likely to respond to a situation in a less logical way [4].

The present paper aims to describe the procedure and results of a replication equal to the Study 6 of Risen and Gilovich in 2008 [1], and propose an extension with certain modifications, inspired by previous replications of the same study under the reproducibility project, which have had contradictory results.

Mathur et al.'s replication [7] of Study 6 showed an effect of cognitive load over the perceived likelihood of being called after both tempting fate and no tempting fate conditions (both reading and not reading scenarios), but no significant main effect for tempting fate (had read vs. had not read) was found. The results could be explained by slight changes that were made to the method. For instance, the sample was made with non-students, and the questionnaire was filled online via Amazon Turk, instead of in a lab setting. Her suggestions for future replications had to do with the possibility to determine if the situation is or not salient to the respondent and verifying the cognitive load on an online setting. Since most of the respondents of her study were non-students it was probable that the situation of being a student in a large class was not perceived as risky and therefore it would not be seen as a tempting fate situation.

To overcome the possible limitation of people not finding the situation salient to them, it was decided to add to the extension two new questions in addition to those asked in the original research, with the aim of identifying if there was a relationship between finding the situation salient and a perceived likelihood of a negative outcome after tempting fate. These new questions were labeled as new variables, Importance and Relevance. Additionally, to surpass the limitation of not being able to control cognitive load in an online environment, the cognitive load method was changed to memorisation.

For the experiment, it was decided to establish two scenarios, a replication of the experiment following the same method proposed by the authors, plus an extension in which the previously stated limitations were addressed. For each situation, two hypotheses were stated:

Replication Hypotheses:

- H0: There is no difference in the perceived likelihood of a negative outcome when presented with TFCL vs. TFNCL (Figure 1).
- Ha: There is a significant difference in the perceived likelihood of a negative outcome when presented with TFCL vs. TFNCL.

For the extension, there is an additional hypothesis, besides the ones mentioned previously:

- H0: There is no correlation between values of Importance or Relevance from the task and perceived likelihood of a negative outcome.
- Ha: There is no correlation between values of Importance or Relevance from the task and perceived likelihood of a negative outcome.

II. Method

The description of the method will be done taking into account the two stages previously mentioned, the replication, which was a strict replication from the method of study 6 from Risen and Gilovich [1] and the extension, which addressed complications regarding how salient participants

would find the situation and the complexity of controlling cognitive load in an online setting.

A. Replication Participants

The replication started with a sample of 40 students from *Universitat Pompeu Fabra* mostly whose primary country of residency was Spain and who were between 20 and 25 years old, compared to the 122 Cornell undergraduates from the original study, who were gathered following a convenient sampling procedure.

In the beginning, the participants were taken to the University Lab Room to have control of loud noises and disturbances. But given the location chosen, many participants would back out because of the effort it meant to move over the lab. For this reason, it was decided to perform some experiments next to the participant's location.

B. Replication Method

In the replication the instructions were given in person and the format used was pen and paper. The step by step started with filling in the consent form, with an addition of a demographic questionnaire for further understanding of the participants. The experiment, just like the one from Risen and Gilovich [1] followed a factorial, between subjects design, the participants were assigned to one of the four conditions (Figure 1).

	Cognitive Load (CL)	**No Cognitive Load (NCL)**
Tempting Fate (TF)	**TFCL**	**TFCNL**
No Tempting Fate (NTF)	**NTFCL**	**NTFNCL**

Fig. 1. Conditions of the experiment.

The methodology from the original study was followed as intended, after completing the consent and demographic form, they were given a printed paper with the situation they had been assigned to, and the correspondent questions. The participants had to situate themselves in a large lecture and imagine how the professor was planning to call on a student because he had no volunteers to answer his/her questions about an assignment. In this scenario, participants either read the assignment for the class or they didn't. They could also be under cognitive load influence or not while reading the situation. Those under load were asked to count backwards by 3 starting with 564.

After finishing reading, the participants had to indicate how likely it was that they would be called by the professor by choosing a number in a Likert scale that went from 0 being not likely to be called at all, and 10 being extremely likely to be called.

To check if participants were indeed under cognitive load they had to report on which number they ended up in, and they also had to indicate how much effort was put on each task, selecting a number between 0 being put all my effort into reading and 6 being put all my effort into counting.

Finally, participants were debriefed and were told about the purpose of the study.

As it was stated before, an extension was proposed due to various factors. First, the complexity of accessing a bigger sample size following a one on one method bearing the resources and time for the experiment, secondly, the poor reliability of a subjective report when controlling cognitive load while in the experiment, and finally, addressing the question of whether the situation of being a student in a large lecture was a proper one to test magical thinking in a non-students sample.

C. Extension Participants

For the extension, both the sampling and the method were made using online sources. A total of 200 participants from all around the world, predominantly Spain and Latin America, mainly between 20 and 24 years old with a completed Bachelors or Master's degree were gathered through a convenient sampling method, using group chats and social media from the researchers.

D. Extension Method

The procedure for the extension is similar to the original study and the replication, nonetheless some differences must be highlighted. Firstly, it was made online, which indicates that both the consent, demographic questionnaire and the survey were completed fully under this modality. Secondly, the method for cognitive load was changed to a memory one, in the extension participants under the CL condition were asked to memorize a pattern (Figure 2) previous to be presented with the situation of the lecture.

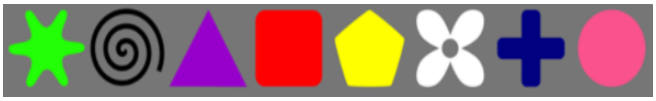

Fig. 2. Pattern to be memorized by the participants under cognitive load.

The last change for the replication was the addition of two questions, with the purpose of finding how salient they found the situation of the lecture and being in the position of a student. The following questions were added at the end of the survey for all four conditions and were named importance and relevance in the analysis of the data:

1. How bad would you feel if you were called on by the professor?
2. If you were a student in the scenario you just read about, how important would it be for you to answer the question correctly in class?

Since the method of cognitive load was changed there was no need to ask about the number the person ended up in neither the effort they have made while counting and therefore those questions were dropped from the extension survey, in exchange, the participant was asked to choose which had been the presented path under a set of four options at the end of the survey.

Both NTFNCL and TFNCL conditions were performed on Google Forms, and for the TFCL and NTFCL conditions a survey was created using Psytoolkit with the aim of limiting

the participants with the possibility of going back to the pattern and reviewing it when asked about it.

The method for the analysis of data was a two way ANOVA for both replication and extension. An additional method was used in the extension, to review if there was any correlation between the importance or relevance given to the situation and the perceived likelihood of a negative outcome.

III. Results

A. Replication

The replication of Study 6 was carried out in person within the premises of the university campus. Participants had to read one of the two self-scenarios, which asked them to imagine themselves in a large lecture and to imagine that the professor is planning to call on a student because nobody has volunteered to answer the question. Two different scenarios were presented to the participants. In the first one, they had done the reading for the class (Not Tempting Fate, or NTF), and in the second scenario, they had not done the reading (Tempting Fate, or TF). Half of the participants who read either scenario were under cognitive load. While reading the story and answering the questions, the participants under cognitive load were asked to count backwards by 3s, starting with 564. Participants specified how likely they believed it was that they would be called on by circling a number from 0 (not at all likely) to 10 (extremely likely). After answering the question, participants under cognitive load were asked to report the number they ended on and specify how much effort they put into each task (counting, answering) by circling a number from 0 (I put all my effort into reading) to 6 (I put all my effort into counting).

There was a total number of 36 participants, equally balanced between the four conditions. We ran a 2 (had not read, Tempting Fate, TF vs. had read, Not Tempting Fate, NTF) X 2 (cognitive load, CL vs. no cognitive load, NCL) ANOVA, but no statistically significant main effect was observed (TF: $F(1, 36) = 1.79$, $p = 0.19$, CL: $F(1, 36) = 2.24$, $p = 0.14$) and no interaction, $F(1, 36) = 0.5$, $p = 0.48$ (Figure 3).

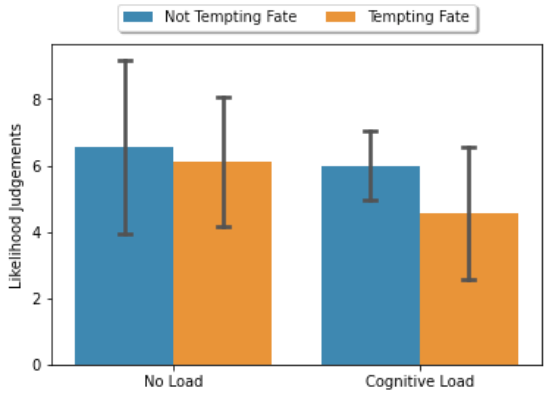

Fig. 3. Subject likelihood judgements in the replication of Study 6. Error bars represent standard deviation.

B. Extension

The extension of the study was carried out online. Participants read the same two self-scenarios as in the replication. Half of the participants who read each scenario were under cognitive load. The cognitive load method was changed from counting to memorising, as participants under cognitive load were asked to memorise a sequence of coloured shapes.

In the extension study, there was a total number of 200 participants, equally balanced between the four conditions. The 2X2 ANOVA was repeated, with no statistically significant main effects (TF: $F(1, 200) = 1.01$, $p = 0.32$, CL: $F(1, 200) = 0.2$, $p = 0.65$) or interaction, $F(1, 200) = 0.03$, $p = 0.86$ (Figure 4).

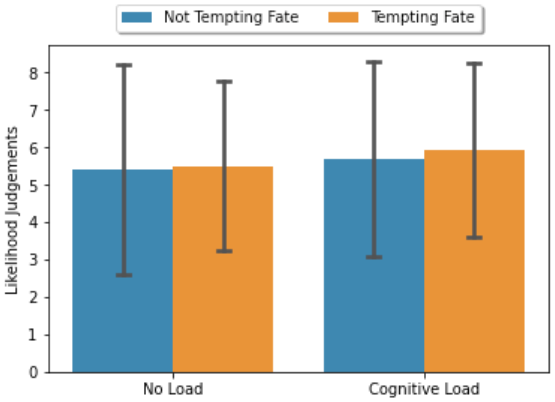

Fig. 4. Subject likelihood judgements in the extension of Study 6. Error bars represent standard deviation.

C. Importance and Relevance

At the end of the extension study, participants were asked to answer two additional questions, by circling a number from 0 (not at all) to 10 (very bad/important):

1. *How bad would you feel if you were called on by the professor?,* quantified as Importance (0-10), and
2. *If you were a student in the scenario you just read about, how important would it be for you to answer the question correctly in class?*, quantified as Relevance (1-10).

The distributions of the two variables were as in Figure 5 and 6. Neither Importance, nor Relevance were normally distributed, as normality tests showed (D'Agostino and Pearson, Importance: $s^2+k^2 = 18.5$, $p < 0.001$, Relevance: $s^2+k^2 = 23.26$, $p < 0.001$). Also, Importance and Relevance were not correlated (Spearman's, $\rho = 0.07$, $p = 0.28$), and could provide two different ways of looking at the data, but each of them is somewhat correlated to the dependent variable (Spearman's, Importance: $\rho = 0.18$, $p = 0.01$, Relevance: $\rho = 0.15$, $p = 0.03$).

D. Selecting by Importance

We selected the participants that answered 4 or 5 to the Importance question. They are participants who would have felt really bad had they been called on. The number of participants who met this criterion is 110.

We ran the same 2X2 ANOVA test and found that there is a statistically significant main effect of cognitive load (CL), $F(1, 110) = 5.49$, $p = 0.02$. We found no other statistically significant effects. An independent t-test confirmed that cognitive load has an effect on the dependent variable, $t(110) = 2.41$, $p = 0.017$ (Figure 9).

We repeated the same process after selecting only participants that answered 4 or 5 to the Relevance question, but found no statistically significant results.

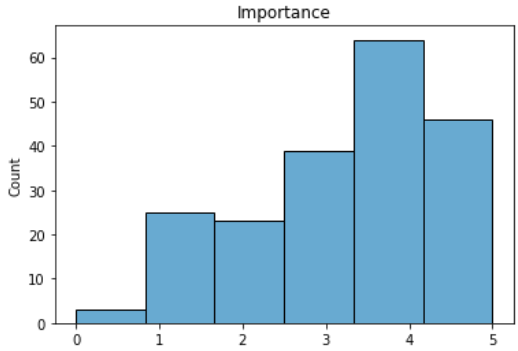

Fig. 5. Distribution of Importance answers in the extension of Study 6. (N=110)

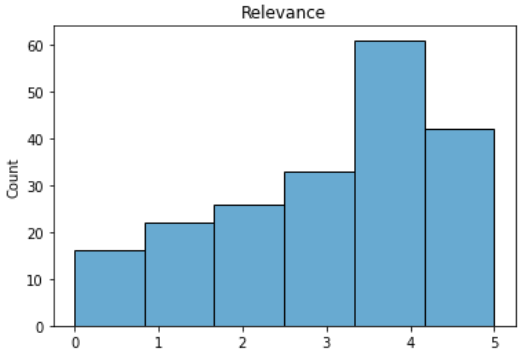

Fig. 6. Distribution of Relevance answers in the extension of Study 6. (N=110)

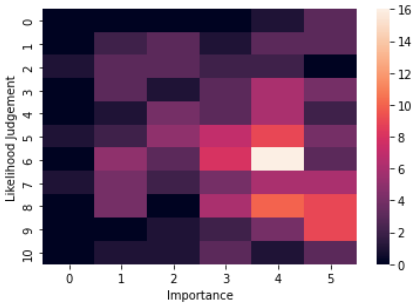

Fig. 7. Relationship between Importance and judgements of likelihood in the extension.

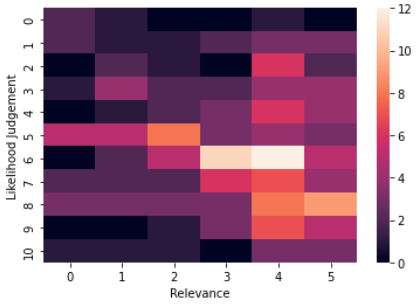

Fig. 8. Relationship between Relevance and judgements of likelihood in the extension.

IV. Discussion

A. What the original experiment showed

Risen and Gilovich (2008) arrived at the conclusion that people did tend to be reluctant to tempt fate, and believed a negative consequence could emerge if they did so, they also showed that there is an enhanced perception of a negative outcome when tempting fate once the person is under the influence of cognitive load.

B. Results Summary

The replication results failed to support the main effect of tempting fate on the likelihood of perceiving a bad outcome as more possible.

The extension results showed that participants who had rated high in Importance (Importance > 3, i.e. those who would have felt bad had they been called by the professor without having read the required reading), when loaded cognitively, had a greater perception of a negative outcome. This is regardless of the condition of tempting fate or not, which is similar to a previous replication done by Mathur et al. in 2020.

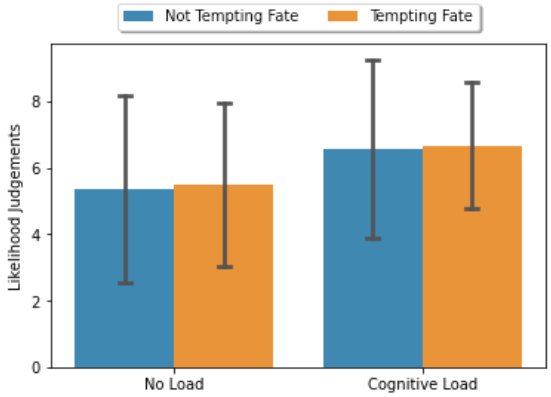

Fig. 9. Subject likelihood judgements in the extension of Study 6, for participants who answered 4 or 5 to the Importance question. Error bars represent standard deviation.

C. Implications

It was not possible to find any effect of tempting fate (TF) on the likelihood of perceiving a bad outcome as more likely. The results might indicate a shift in the conception of what the consequence would be after not reading for the class from 2008 to 2021, and could suggest that those answering the survey for the present study live under a different educational culture than those surveyed in the first study(Cornell University Undergraduates), similarly, it could mean that not reading for a lecture is not perceived as an act of tempting fate nowadays and therefore is not an appropriate situation to use on an experimental design to measure the belief in tempting fate and magical thinking.

However, after selecting only those participants who had answered either 4 or 5 to the question *How bad would you feel if you were called on by the professor?,* a statistically significant effect of cognitive load on the dependent variable was found, suggesting that indeed cognitive load and therefore the dual thinking perspective (System 1 and System 2) might be involved in the way people make judgements about a situation, and magical thinking.

If that is the case, then the difference between NCL and CL could be explained by what Risen [3] refers to as the

impossibility of System 2 to correct the irrational judgements made by System 1, once its capacity is occupied with another task.

The present study could not explore if the ability and motivation of being rational played a role in the process of deciding if a person was going to be called on or not. Nevertheless, it is relevant to highlight that some participants while in the replication displayed some of these traits when asked how likely they were to be called by the professor by asking questions such as: *How big exactly is the room?* or *What was the number of students in the room?*, showing a slight tendency to seek for more information, probably to make a more thoughtful or informed decision.

It could be interesting to explore how the ability and motivation of being rational play a role in how people judge a situation related to tempting fate and add it to further replications of the original study.

D. Comparison with previous replications

Maya Mathur et al. [7] conducted a high-powered (four university sites, $N = 754$), pre registered replication focused primarily on the interaction effect between cognitive load and tempting fate. Similar to the replication performed in the present paper, they failed to reject the null hypothesis and prove the main effect of tempting fate or the interaction, obtaining very similar results (Figure 10 and Figure 11). It's worth noting that the means for likelihood are lower in their replication.

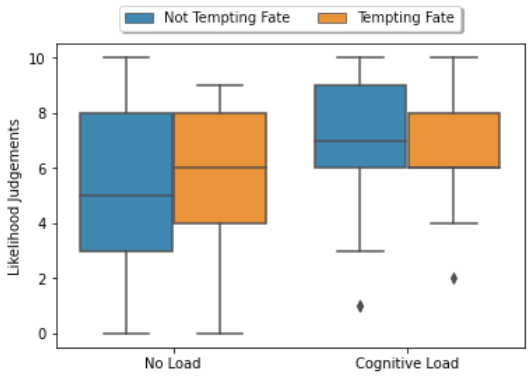

Fig. 10. Subject likelihood judgements in the extension of Study 6.

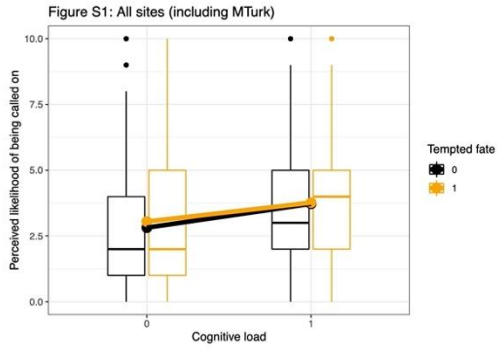

Fig. 11. Subject likelihood judgements in Maya Mathur et al.'s 2020 replication of Study 6.

E. Limitations

Some limitations were found while developing the experiment. First, the lack of correct measure of cognitive load in both lab and online settings. Additionally, the fact that non-student participants and even some student participants seemed not to find the situation of the lecture salient to them, which leads to vague results if the objective is to measure magical thinking when tempting fate. Finally, the authors from the original study assumed that higher responses of a perceived likelihood of a negative outcome after tempting fate are a consequence of the belief in tempting fate, which could be explained by other factors. A solution to this assumption could be asking the participants why they responded the way they did to confirm if high rates in likelihood of being called are a result of believing in tempting fate.

V. Conclusion

Risen and Gilovich (2008) arrived at the conclusion that people did tend to be reluctant to tempt fate, and believed a negative consequence could emerge if they did so. The replication of the study performed by the authors of the current study failed to replicate the findings of the original study.

A further methodological extension was proposed, with three modifications: (1) two new variables were added (Importance and Relevance) to control if whether the participants cared or not about the tempting fate scenario of the student, (2) changes in sampling, and (3) the method by which participants were loaded cognitively, which will be described in detail in the method section.

The results showed that participants who had rated high in Importance (i.e. those who would have felt bad had they been called by the professor without having read the required reading), when loaded cognitively, had a greater perception of a negative outcome. This is regardless of the condition of tempting fate or not, which is similar to a replication done by Maya Matur in 2016. The other main effect from the original study (tempting fate has an effect on perception of negative outcomes) and the interaction between cognitive load and tempting fate could not be replicated.

As demonstrated in the results, the situation of the students is not ideal to measure magical beliefs among people, because it assumes the situation is perceived as risky by everyone, when the results from the extension seen above confirm it is not. In this case, it could be interesting to explore which are the variables, either education, cultural beliefs, race, ethnicity or any other, that could be related to how subjects consider a situation as tempting fate or not.

Finally, the present study could not replicate the original null hypothesis of the paper, that "tempting fate" increases the perception of a negative outcome as being more likely, or that cognitive load amplifies that effect.

References

[1] Risen, J., Gilovich, T, "Why People Are Reluctant to Tempt Fate", Journal of Personality and Social Psychology, 95, pp. 293-307, 2008.

[2] Zusne, L., Jones, W. H., Anomalistic Psychology: A Study of Magical Thinking, 1989.

[3] Risen, J., "Believing What We Do Not Believe: Acquiescence to Superstitious Beliefs and Other Powerful Intuitions", Psychological Review, 123, 2015.

[4] Kahneman, D., Thinking, Fast and Slow (1st ed.). Farrar, Straus and Giroux, 2013.

[5] Callan, Mitchell & Sutton, Robbie & Harvey, Annelie & Dawtry, Rael, “Immanent justice reasoning: Theory, research, and current directions.”, Advances in Experimental Social Psychology, 49, pp. 105-161, 2014.

[6] McKean, E. (Ed.) The New Oxford American dictionary (2nd ed.). New York: Oxford University Press, 2015.

[7] Mathur, M., et al., “Many Labs 5: Registered Multisite Replication of the Tempting-Fate Effects in Risen and Gilovich (2008)”, Advances in Methods and Practices in Psychological Science, 1–11, 2020.