Graduate thesis

NTNU
Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Psychology

10153

Revisiting Temporality Effect: Replication and extension of the tendency in a coin toss scenario, and in good outcomes

Graduate thesis in Psychology
Supervisor: Subramanya Prasad Chandrashekar
May 2023

NTNU
Norwegian University of
Science and Technology

10153

Revisiting Temporality Effect: Replication and extension of the tendency in a coin toss scenario, and in good outcomes

Graduate thesis in Psychology
Supervisor: Subramanya Prasad Chandrashekar
May 2023

Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Psychology

NTNU
Norwegian University of
Science and Technology

Preface

This study exist as part of an empirical replication project, planned and supervised by Subramanya Prasad Chandrashekar, the supervisor for this bachelor group. The bachelor project is part of the spring 2023 PSY2900 course, of NTNU. The topic and initial literature was presented early in the year by the supervisor. The exact experiment of the chosen target was chosen based on a few options picked out by the supervisor. The hypotheses were formulated from the original target replication studies, with supervision from the supervisor. The ideas for the extension were primarily my own, but were based on a general discussions of the topic during meetings with the bachelor group. The additional hypotheses were derived from a discussion with the supervisor.

Ethical approvals was handled by the supervisor. The translation of the first study was as a collaborate effort among all students of the bachelor group, as was the recruitment of participants. Recruitment of participants of the other studies was handled by the supervisor. The pre-registrations and surveys was performed in collaboration with the group, and the supervisor. The translation of the survey into data material was done by the supervisor, but analyses was work of my own. Literature search beyond the initial, was mainly conducted using Apa Psycnet, Oria, and Google Scholar, and was work of my own. For references APA 7th was used. Interpretations of the results are my own. Throughout the semester we had the possibility to get feedback from the supervisor, and several class meetings was conducted.

I wish to thank Subramanya Prasad Chandrashekar, for introducing me to the topic, and for valuable feedback throughout the whole process. I also wish to thank the other students in the bachelor group, for useful discussion and support throughout the project. I also wish to especially thank my girlfriend, for her constant support.

Abstract

The temporality effect is the tendency to mutate the latest event, in an independent sequence of events (Byrne et al., 2000). We aimed to revisit the temporality effect, with two pre-registered replications, and an exploratory extension. We conducted a replication of the coin toss scenario in Miller and Gunasegaram (1990), and successfully replicated the original findings of a temporality effect. This was consistent on mutation, in addition to attribution of guilt and blame. We also conducted a replication of the temporality effect in good outcomes (Byrne et al., 2000), and successfully replicated the findings on mutation, but not on relief. Our findings did not show a difference in the two groups in relation to relief. Our extension found an elimination, and a reversal, of the standard temporality effect, implying that choice might have an effect on such sequence events. Generally, the findings arguably support the counterfactual model, to a larger extent, than the crediting causality model.

Revisiting Temporality Effect: Replication and extension of the tendency in a coin toss scenario, and in good outcomes

Temporality effect is the tendency to mutate the last event in an independent sequence of events, and is one of the factors affecting mutability in counterfactual thinking. The effect has also been linked to causal and emotional attribution, such as guilt, blame, and relief (Miller & Gunasegaram, 1990; Byrne et al., 2000). However, there have been some differences in findings of such emotional attribution, especially in relation to mutability. Moreover, earlier studies have hypothesized different theoretical mechanisms explaining the tendency, two of these being; the crediting causality model (Spellman, 1997), and the counterfactual model (Byrne et al., 2000). In addition, a general importance is put on conducting replications, especially of earlier psychological studies. This thesis therefore aim to revisit the findings in temporality effect, attempting to closely replicate the coin toss scenario in Miller and Gunasegaram (1990), and the good outcome experiment in Byrne et al., (2000). We also aim to interpret the results within the counterfactual model, as well as the crediting causality model, aiming to discuss their applicability. In addition, we also conduct an extension, examining the possible effects of implementing choice in a temporal order sequence.

Counterfactual Thinking

Counterfactual thinking is the tendency to imagine alternatives to reality, reflecting on past events. When executing counterfactual thinking, humans mentally change certain elements of an event, and imagine the supposed outcome these changes would ensue. Counterfactual thinking appears in humans everyday life, and holds implications on emotional and cognitive processes (Segura et al., 2002). Furthermore, counterfactual thinking helps

people learn from the past, thus preparing them for the future (Byrne, 2016). An everyday example could be missing the bus to class one day. By thinking "if only I had left five minutes earlier, I would have got to the bus in time", one could learn for future days with class to always leave five minutes earlier. Several factors affect how counterfactuals are produced for events, even though every event has, in theory, limitless possible alternatives (Walsh & Byrne, 2004). These factors affect the amount, and the availability, of counterfactuals produced. Kahneman and Miller (1986) introduced some of these factors, and hypothesized that the more available the counterfactuals of a feature are, the more mutable the feature is. Furthermore, they hypothesized that the more mutable a feature is, the more it is attributed causality. By connecting this with the tendency of victim-blaming in cases of violence, they portraited part of the importance of identifying these factors. Reasoning that in the case of victim-blaming, the tendency to attribute responsibility and blame to the victim derived from how the victims actions were easier to mutate (Kahneman & Miller, 1986). The factor this thesis will focus on is the temporal order of events, also called the temporality effect.

Temporality Effect

The temporality effect is the tendency to mutate the last occurrent event in a sequence of independent events (Walsh & Byrne, 2004). To our knowledge, the first study on the effect was by Miller and Gunasegaram (1990). Their article presented a simple coin toss scenario where two people flipped a coin in order, and if they got the same outcome they would both win $1000. The scenario described a losing outcome, as the first person (Jones) flipped heads, and the second person (Cooper) flipped tails. When participants were asked about who they thought would feel more guilt, and be blamed more by the other, a large majority of participants answered Cooper. The majority also answered that "Cooper tossing a head" was more readily in mind, when asked to imagine how they could have won (Miller &

Gunasegaram, 1990). In actuality however, the outcome is purely based on chance, and neither Jones or Cooper should be given more causality for the outcome. The findings thus demonstrated how the temporality effect affects peoples judgements in sequences of two independent events. The effect is also relevant in everyday situations, actively effecting peoples judgements. For instance, students perceived test questions more fair if they were set before the student finished studying, opposed to after (Miller & Gunasegaram, 1990). People also judge the last game in a team's season to be the most important (Walsh & Byrne, 2004), however, each game throughout the season is equally important for the team's final placement. In addition, temporality effect occurs in both good and bad outcomes, and is based on the order of the events actual occurrences, rather than their order of mention (Byrne et al., 2000).

As previously mention, earlier studies have argued for the existence of a link between mutability and causal attribution, and between causal attribution and emotional attribution (Kahneman & Miller, 1986; Miller & Gunasegaram, 1990). Causal attribution reflects how much causal weight an event is given, especially related to the occurred outcome. The coin toss exemplifies this possible link, as Cooper`s actions are viewed as more mutable, and he is attributed more guilt and blame for their loss (Miller & Gunasegaram, 1990). However, later studies have found a distinction between mutability and emotional attribution, as guilt and blame were attributed to the last event, even though the first event was more mutated. The attribution of emotions thus seemed to rather exhibit the standard temporality effect, following the representation of the factual events of a scenario (Walsh & Byrne, 2004). Either way, the temporality effect plays a role in humans judgement about emotional and causal attribution, and affects everyday situations.

Theoretical Mechanisms

One of the hypotheses that introduced theoretical explanations for the temporality effect, was Spellman`s (1997) crediting causality hypothesis. The model explains five steps in which causal attribution of an outcome is based on. The causal attribution is based on calculations on which event, in the outcome sequence, most changed the probability of the occurred outcome. The calculations occur before the first event, then after each event, and the calculation of the second event`s effect is based on the outcome of the first event. After the events, the change in probability from each event is compared. Essentially, the event in the sequence that had the biggest probabilistic impact on the occurred outcome, is attributed with the most causality (Spellman, 1997).

The coin toss scenario (Miller & Gunasegaram, 1990) can provide an exemplification of the model in an independent two event sequence, as it is a 50/50 chance scenario. Before the first throw the probability of a winning outcome is 50%; they either get the same coin or not. This continues after the Jones flips, as the probability of flipping heads/tails is the same each flip. Thus, Jones changed 0% of the probability of the losing outcome, as it was 50% both before and after his event. However, when Cooper flips, the probability of losing changes to either 0% or 100%. This because he either flipped different, resulting in a 100% probability of losing, or he flipped the same, resulting in a 100% probability of winning. As the outcome is that they lost, the causality of the loss is attributed to Cooper, since his flip changed the probability of the losing outcome the most.

A second theory that provided a model for the theoretical mechanisms, was the counterfactual model by Walsh and Byrne (2004). The counterfactual model presented that there were six tenets that the mechanisms behind temporality effect were based on, regarding the representation and mutation of the counterfactual possibilities imagined. The first three

tenets describes that people keep in mind only parts of a few, true counterfactual possibilities, guided by the winning conditions of the event. This is because of limited space for storing information in the working memory, thus having to prioritize imagining only a few, effective possibilities. As counterfactuals are only effective if they undo the actual outcome, the working memory uses its limited space for counterfactuals that reach a winning outcome. The fourth tenet is that people mutate facts so they match with the counterfactuals of the winning conditions. The fifth tenet is the presupposition of the first event of the facts, making it the element that people try to match with the winning counterfactual possibilities. This presupposition derives from how humans create mental models of incoming information, to coherently understand the facts of the situation. The first event of an occurred sequence, possibly because of humans perception of time as linear, acts therefore as the anchor and initialization of the model, that all following factual events are merged into (Byrne et al., 2000). This presupposition of the first event is theorized to be the reason for the greater mutability of the last event, resulting in the temporality effect (Kahneman & Miller, 1990; Byrne et al., 2000). Thus, the temporality effect does not occur primarily because of a wish to mutate the last event, but an opposition to mutate the first event. The last event is more mutable in comparison to, and because of, the first event, and not principally because of some attribute to the last event itself.

The standard temporality effect appears if the presupposed element of the fact, manages to match with the explicitly held winning condition. However, if the first factual element does not match, it is the one that is mutated. If the first element is mutated, a reversal of the temporality effect occurs. This reversal relates to the sixth tenet; that some elements in a possibility are kept more explicitly in mind, while other parts of the same possibility are only kept implicitly. Which elements of the same possibility people keep explicitly or implicitly can be effected by how the winning conditions were descripted (Walsh & Byrne,

2004). In general, in an independent losing event sequence, people create a model of the facts, what actually occurred. In addition, people create a model of counterfactual alternatives. This model contains only the possible winning outcomes, which are based on the depicted winning conditions. People then try to match the factual model with the counterfactual model, and generate a counterfactual alterative constructing a complete match (Walsh & Byrne, 2001).

An important point to make when discussing the earlier frameworks, is how Byrne et al., (2000) findings contradicted Spellman`s (1997) crediting causality hypothesis. Their experiment presented a scenario where two individuals would win $1000 each, if they drew the same colored card from a deck (see method section for more descriptive details). The participants were assigned to either a same card, or the different card scenario. In both scenarios a technical hitch appears after the first person draws, making the first participant having to draw again. In the same card scenario the pre- and post-hitch color is the same, and in the different card scenario the pre- and post-hitch color is different. The results showed the standard temporality effect in the same card scenario, and an elimination of the effect in the different card scenario. The findings showed that making a counterfactual readily available for the first event, eliminated the temporality effect. When creating a counterfactual model, people equally often chose to base it on the existing facts, as on the presupposition of the first event. Thus, the availability of a counterfactual countered the presupposition of the first event, making people equally choose the first or second person. The technical hitch thus rejected the crediting causality hypothesis, because if people actually updated their calculations after each event, there should be no difference between the same- and different card scenario (Byrne et al., 2000).

Need For Replication

For over a decade now there has been an increase in focus on replicability and robustness of findings, and a general understanding of an ongoing replication crisis (Schrout & Rodgers, 2018). One reason for this it the findings from a large replication study, that discovered that of 100 earlier published studies, only 36% of the replications found a significant result (Brandt et al., 2014). The discussion of the need for replication is, among other things, centered around false-positives and the possible ways to publish unreliable results, for instance with "*p*-hacking" (Vadillo et al, 2016). Arguably, the consequences of false positives in psychological studies can have a negative effect beyond the scientific field. One example may be the relation between psychology and the legal system (Ogloff & Finkelman, 1999). In an attempt to counter this ongoing crisis, open science practices have been more implemented in current research. Open science is the motion to make scientific practices and reports more transparent and accessible (Hagger, 2019), for instance by conducting a preregistration.

As several of the studies on temporality effect were published over 15 years ago, we wanted to revisit the effect, employing the newer methods and ideals of open science.

Choice of replication articles

We chose the coin toss scenario in Miller and Gunasegaram (1990), and the good outcome scenario in Byrne et al., (2000), as our targets of replication. The choice of the two articles was mainly based on their big influence on the topic, and discussions, around temporality effect. However, we also wanted to increase the sample sizes, and calculate and compare effect sizes.

The coin toss scenario in Miller and Gunasegaram (1990) has a simple design, and is used as an introduction and exemplification of temporality effect in several articles (Byrne et

al., 2000; Walsh & Byrne, 2004; Segura et al., 2002). As previously mentioned the article was also the first, to our knowledge, to explicitly study the temporality effect, and has since been cited over 330 times. Their article also demonstrated how the temporality effect appears in everyday situations, further providing relevancy of the topic. In the coin toss scenario they found a clear demonstration of the tendency to rather mutate, and attribute guilt and blame, to the last person in an independent sequence of two events. The study is therefore a fundamental contribution in exploring the temporality effect, and the connection between mutability and causal attributions.

The study by Byrne et al., (2000) consisted of five experiments, all exploring the temporality effect in different scenarios. One of the important findings in the article was that including a technical hitch had a mediating effect on the temporality effect. As previously mentioned, this technical hitch ruled out some other theories, such as the crediting causality hypothesis. The article also presented the abovementioned counterfactual model theory, which proposed an explanation of the underlying mechanisms behind temporality effect. In general, the study arguably contributed with several important findings and interpretations. Our replication focuses on experiment three in their article, where they test for the existence of temporality effect in good outcomes. They argue that just as counterfactuals relates to learning from earlier mistakes, they may relate to learning from earlier success as well. Knowledge is gained by having a successful experience, which can be used to increase the chance to repeat the actions leading to the outcome. The experiment found that the temporality effect exists in good outcomes as well. In addition, an elimination of the temporality effect appears when including the technical hitch.

The present investigation`s aim is to try to closely replicate both these experiments findings, following LeBel et al., (2019) criteria of replication (see supplementary materials for comparison).

Extension

In addition to the two replications, we also added an extension, implementing choice in a good outcome scenario. We based the scenario on the same scenario as the target replication, namely the experiment three in Byrne et al., (2000). The original experiment made the first individual draw again because of a technical hitch, an element that was out of the participant control. However, we rather explained that the first participant had the choice to either draw again or not, where she chose to draw again in both scenarios. Several studies have presented how choice and actions are of importance to mutability and counterfactuals (Girotto et al., 1991; Byrne & McEleney, 2000), which we tried to incorporate in our extension. However, we hypothesized that participants would choose the last person in all three questions, across both scenarios.

Open Science Practices

We preregistered all the experiments on the Open Science Framework (OSF), before data collection. Power analyses, pre-registration, and all materials are available in the supplementary material. All measures, manipulations, and exclusions are reported, and data collection was completed before analyses. Pre-registration is available at: https://osf.io/qkvb5 (study 1), https://osf.io/k5tm7 (study 2), https://osf.io/xjwq7/ (study 3). All the surveys were digital and first carried a brief description about the study, and sought consent.

Method

Study 1- Replication of Miller and Gunasegaram (1990)

Participants

There was a total of 211 participants recruited (138 females, 72 males, 1 other), with age ranging from 19-90 years (M = 28.5, SD = 13.68). All the participants were Norwegian speaking adults, informed that they would remain anonymous.

Procedure

The participants were recruited mostly through convenience sampling, with the student group sending the survey link to friends and family. There were also printed out flyers containing a QR code leading to the digital survey. These flyers were handed directly out to students on Dragvoll Campus, in addition to leaving the flyers around campus for anyone to scan and participate. The survey included a replication and an exploratory scenario based on the coin toss scenario, however this thesis will focus only on the replication scenario. Previous to the scenario there was a paragraph asking the participants to try to form a detailed understanding of the decision-makers involved. After reading the scenario they were asked three questions, and asked to choose one of the two possible answers (Jonas or Kristian) to each of the questions. The questions were designed to report mutation, and attribution of guilt and blame. (For full details, see pre-registration in supplementary materials).

Results

We conducted three one-sample Z-tests to test the three hypotheses. Participants showed a tendency to mutate the latest event (76.3%), more than the first event (23.7%; binomial z = 7.64, p < .001; Cohen`s g = 0.26, 95% CI [0.20, 0.32]). The participants also showed a tendency to attribute guilt to the second player (93.4%), rather than the first player (6.6%; binomial $z$ = 12.6, $p$ < .001; Cohen`s g = 0.43, 95% CI [0.40 , 0.47]). Lastly, participants also showed a tendency to assume that the first player would blame the second

player more (91.9%), rather than the second player blaming the first player more (8.1%; binomial $z = 12.19$, $p < .001$; Cohen`s g = 0.42, 95% CI [0.38 , 0.46]).

Table 1

Mutation, Guilt and Blame statistics for study 1

Dependent variables	Participants	%
Mutation		
First	50	24
Second	161	76
Guilt		
First	14	7
Second	197	93
Blame		
First	17	8
Second	194	92
N	211	

Study 2- Replication of Byrne et al., (2000)

Participants

There was a total of 339 participants (167 males, 160 females, 10 non-binary, 2 prefer not to say), with age ranging from 10-77 years ($M = 38.5$, $SD = 13.9$). The American participants were recruited based on location, and recruited online through the Prolific platform (Prolific.com). Based on our exclusion criteria and Prolific own criteria's, we believe

there was a mistake done with the one participant answering their age as 10. The lowest age is therefore in actuality 18 years.

Procedure

The survey presented an explanation of the task, including instructions to read the scenario carefully, to complete the questions in the order presented, and refrain from changing an answer after writing it down. The participants were randomly assigned to read either the same card scenario (n = 168), or the different card scenario (n = 171). The scenarios and questions were designed to replicate the third experiment in the original study by Byrne et al., (2000), and therefore mostly followed the same descriptions.

The general description before both the scenarios read as follows: “Imagine two individuals (Jones and Brady) who take part in a television game show, on which they are offered the following very attractive proposition. Each individual is given a shuffled deck of cards, and each one picks a card from his own deck. If the two cards they pick are of the same color (i.e., both from black suits or both from red suits), each individual wins £1,000. However, if the two cards are not the same color, neither individual wins anything. “

In the same card condition the following description of the technical hitch is then presented: “Jones goes first and picks a red card from his deck. At this point, the game show host has to stop the game because of a technical difficulty. After a few minutes, the technical problem is solved and the game can be restarted. Jones goes first again, and this time he draws a red card. Brady goes next and the card that he draws is also a red card. Thus, the outcome is that both individuals wins £1,000 each.” The different card condition description of the technical hitch is almost identical, only differing as Jones first draws black and after the hitch red, and then Brady draws a red card as well.

After reading one of the two scenarios, they were asked to answer three questions regarding the scenario, having to choose either “Jones” or “Brady” in each question. The first question aimed to check for which persons actions were easiest to mutate. The next question asked who they predicted would experience more relief. We also added a question on who they would predict would feel more responsible for winning, and this questions was added in the extension as well. The survey then asked for age and gender.

Notably, some alterations from the original study were implemented, one alteration being that both questions were binary. This excluded the option “neither” in relief, and the sentence completion task of mutation. This alteration was done because coding text would be impossible considering the time and resource restraint of this thesis. Another alteration was the addition of one question, this being the question about “Who would you predict would feel more responsible for the winning outcome?, which was added as an exploratory question.

Results

We conducted one-binomial tests for all three hypotheses in the same card scenario. In the mutation question, participants mutated the second event (70.2%), more than the first (29.8%; binomial $n = 168$, $k = 118$, $p < .001$; Cohen`s $g = 0.20$, 95% CI [0.13, 0.27]). The participants also showed a tendency to assume the second person would feel more relief (59.5%), rather than the first (40.5%; binomial $n = 168$, $k = 100$, $p = .016$; Cohen`s $g = 0.10$, 95% CI [0.02, 0.17]). This difference was statistically significant, however the effect size was small. In addition, the participants attributed more responsibility to the second person (69.6%), rather than the first (30.4%; binomial $n = 168$, $k = 117$, $p < .001$; Cohen`s $g = 0.20$, 95% CI [0.13, 0.27]).

We conducted one binomial tests across each hypotheses in the different card scenario as well. There were no tendency for participants to mutate the second (53.2%), rather than the

first event (46.8%; binomial n = 171, k = 91, p = 0.444; Cohen`s $g$ = 0.03, 95% CI [-0.04, 0.11]). In the question with relief, participants attributed more relief to the first person (57.9%), rather than to the second person (42.1%; binomial $n$ = 171, $k$ = 72, $p$ = 0.046; Cohen`s g = -0.08, 95% CI [-0.15, -0.004]). This difference was statistically significant, however the effect size was small. For responsibility, the participants showed no tendency to attribute more to the second person (56.1%), rather than the first (43.9%; binomial n = 171, k = 96, p = 0.125; Cohen`s g = 0.06, 95% CI [-0.01, 0.14]).

We also conducted a two-proportions Z-test testing for a higher rate of mutation in the same card scenario compared with the different card scenario. Results showed that there was a significant difference between the two groups, with people mutating the second event at a higher rate in the same card condition, compared with the first (Z = 3.22, k1 = 118, n1 = 168, k2 = 91, n2 =171, p < .001; Cohen`s h = 0.35, 95% CI [0.14, 0.57])

Table 2

Mutation, relief and responsibility statistics for study 2

	Scenario 1 (same card)		Scenario 2 (different card)	
Dependent variables	Participants	%	Participants	%
Mutation				
First	50	30	80	47
Second	118	70	91	53
Relief				
First	68	40.5	99	58
Second	100	59.5	72	42
Responsible				
First	51	30	75	44

Second	117	70	96	56
n	168		171	

Study 3- extension of temporality effect

Participants

There was are total of 324 participants (179 males, 140 females, 5 non-binary), with age ranging from 18-77 years (M = 39.5, SD = 13.6). The American participants were recruited based on location, and recruited online through the Prolific platform (Prolific.com).

Procedure

The survey first presented an explanation of the task, including instructions to read the scenario carefully, to complete the questions in the order presented, and refrain from changing an answer after writing it down. The participants were randomly selected to read either the same marble scenario (n =158) or the different marble scenario (n = 166). The scenarios and questions were based on the third experiment from Byrne et al., but altered parts to check for an possible effect of choice implementation.

The general description in both scenarios before the technical hitch read as follows: "Imagine two individuals, Anne and Joan, who are offered the following proposition. Each individual is given a sack of marbles, and each one draws a marble from her own sack. If the two marbles they draw are of the same color (i.e., both are blue or both are white), each individual wins £1,000. However, if the two marbles are not the same color, neither individual wins anything."

The description in the same marble scenario:

“Anne goes first and picks a blue marble from her sack. At this point Anne gets the option to choose to either draw again, or to not draw again and keep her blue marble. She chooses to draw again and draws a blue marble. Joan goes next and the marble she draws is a blue marble. Thus, the outcome is that both individuals wins £1,000 each.”

The different marble scenario altered the description by presenting that Anne draws white and then blue, and then Joan draws blue as well.

The three questions then asked were identical to the ones asked in the replication study, regarding mutation, relief, and responsibility. However, the alternatives given to question 1 about mutation had slightly different alternatives in the different scenarios. With the same card scenario the alternatives being:

- Anne drew a white marble the second time
- Joan drew a white marble

Whereas in the different marble scenario, the alternatives was presented as such:

- Anne chose to not draw again
- Joan drew a white marble

There are several alternative scenarios possible when adding the option of choice (for instance that both participants got to choose), however we concluded with this presentation of options being the best suited for our study.

Results

We conducted one binomial tests for the three hypotheses across both scenarios. For the same card scenario the participants showed no tendency to mutate the first (56.3%), rather than the second (43.7%; binomial n = 158, k = 69, p = 0.130; Cohen`s g = -0.06, 95% CI [-0.14, 0.02]). The participants showed a tendency to attribute more relief to the first person

(60.1%), rather than the second person (39.9%; binomial n = 158, k = 63, p = 0.013; Cohen`s $g$ = -0.10, 95% CI [-0.18, -0.02]). For responsibility, the participants showed no tendency to attribute more to the second person (53.8%), rather than the first (46.2%; binomial $n$ = 158, $k$ = 85, $p$ = 0.382; Cohen`s g = 0.04, 95% CI [-0.04, 0.12]).

For the different card scenario the participants showed a tendency to mutate the first event (72.3%), rather than the second (27.7%; binomial n = 166, k = 46, p < .001; Cohen`s $g$ = -0.22, 95% CI [-0.29, -0.15]). In addition, the participants showed a tendency to attribute relief to the first person (77.7%), rather than the second (22.3%; binomial $n$ = 166, $k$ = 37, $p$ < .001; Cohen`s g = -0.28, 95% CI [-0.34, -0.21]). The participants also showed a tendency to attribute responsibility to the first person (78.9%), rather than the second (21.1%; binomial n = 166, k = 35, p < .001; Cohen`s g = -0.29, 95% CI [-0.35, -0.23]).

Table 3

Mutation, relief and responsibility statistics for study 3

	Scenario 1 (same marble)		Scenario 2 (different marble)	
Dependent variables	Participants	%	Participants	%
Mutation				
First	89	56	120	72
Second	69	44	46	28
Relief				
First	95	60	129	78
Second	63	40	37	22
Responsible				
First	73	46	131	79

Second	85	54	35	21
n	158		166	

Summary of Replication Findings

We followed Lebel`s criteria (2019) for replication as our framework for interpreting the results. A complete comparison for both studies is presented in table 4.

In our first study of the coin toss scenario, the replication findings support the original findings by Miller and Gunasegaram (1990). These findings are consistent across all three hypotheses, on mutation, guilt and blame attribution. The replication thus support that people show a tendency to rather mutate, and attribute guilt and blame, to the last event in an independent sequence of two events.

Our second study, on experiment three in Byrne et al., (2000) successfully replicated the original findings on mutation. People showed a tendency to rather mutate the second event in the same card scenario, while the tendency was eliminated in the different card scenario. The same was apparent in the question we added of responsibility. These findings support that the standard temporality effect, and its elimination, appears in good outcomes as well. However, the results was somewhat inconsistent in the question regarding relief. Following Lebel`s criteria for replication (2019), our study successfully replicated the original findings. However, the effect sizes were below small in both studies (based on the benchmark suggested by Cohen, 1977). Furthermore, when qualitatively considering the participants and percentages, the between group difference was arguably near none. However, the original study reported a significant difference in between groups comparison. Some possible explanations for these findings will be further discussed in the general discussion section.

Table 4

Comparison and interpretation of original studies and replications

Study	Original study effect size	Replication study effect size	Interpretation
Study 1, Miller and Gunasegaram (1990)			
	Cohen`s *g*	Cohen`s *g*	
Mutation	0.39 [0.32, 0.45]	0.26 [0.20 , 0.32]	Signal- inconsistent, smaller
Guilt	0.36 [0.29, 0.44]	0.43 [0.40 , 0.47]	Signal- inconsistent, larger
Blame	0.42 [0.36, 0.48]	0.42 [0.38 , 0.46]	Signal- consistent
Study 2, experiment 3 in Byrne et al., (2000)			
Scenario 1 (same card)	Cohen`s *g*	Cohen`s *g*	
Mutation	0.19 [0.04, 0.34]	0.20 [0.13, 0,27]	Signal- consistent
Relief	0.13 [-0.02, 0.28]	0.10 [0.02, 0.17]	Signal- consistent
Scenario 2 (different card)			
Mutation	-0.03 [-0.21, 0.15]	0.03 [-0.04, 0.11]	No signal- consistent
Relief	-0.01 [-0.17, 0.15]	-0.08 [-0.15, -0.004]	Signal- consistent
Across conditions	Cohen`s *h*	Cohen`s *h*	

Mutation	0.45 [-0.01, 0.91]	0.35 [0.14, 0.57]	Signal- consistent

Summary findings of extension

We hypothesized that the standard temporality effect would appear in both scenarios, across all three questions. However, in the same marble scenario the results show no evidence for a difference between the first and second person in mutation, or responsibility. For relief the results showed a tendency for people to rather choose the first person, and even though it was significant, the effect size was below small. We therefore interpret that there is no difference in attribution of relief either. These findings reveal no clear temporality effect in the same marble scenario regarding all three questions. This elimination of the standard temporality effect contradicts our hypotheses, and the findings seem to indicate that implementing choice has an effect on temporality effect in independent two event sequences.

Moreover, in all three conditions in the different marble scenario we found a reversal of the standard temporality effect, as people showed a tendency to rather choose the first event. The results further shows a difference between the two scenarios, with the opposing findings of an elimination, and a reversal of the temporality effect. The results therefore generally contradicts our hypotheses, and yields further support for the notion that implementing choice has an effect on temporality effect in independent two event sequences.

General Discussion

This thesis revisited the temporality effect by conducting two close, pre-registrated replications of the temporality effect, one for the coin toss scenario in Miller and Gunasegaram (1990), and one for the good outcome scenarios in Byrne et al., (2000). The

replications had larger sample sizes than the original studies, with an increase from 88 to 211 in study 1, and from 94 to 339 participants in study 2. We successfully replicated the findings from Miller and Gunasegaram (1990), as participants showed a tendency to mutate, and attribute guilt and blame, to the last event in an independent two events sequence. The replication effort yields support for the original findings of a temporality effect, furthering the empirical strength of the effect. It seems as the coin toss scenario is effective in making participants execute the temporality effect, arguably partly because of its simplicity in design. Moreover, our study was conducted with a Norwegian sample, arguably strengthening the generalizability of the findings.

Looking at these results in terms of the two theoretical mechanisms previously discussed, both explanations are applicable. In the counterfactual model the first element in the factual model would match with the first element in the counterfactual winning model; Jones tosses head. As there is an instant match this counterfactual is chosen, which includes that Kristian tosses head. Furthermore, there are no competing tendencies of mutability, making the presupposition of the first event the main tendency. In addition, the crediting causality explanation also matches with the findings. As Kristian is the one that causes the most probabilistic change to the outcome, his action are seen as most causal. In general, the coin toss was successfully replicated, but could not provide further distinction between the mechanical theories.

We also successfully replicated the findings on mutation in Byrne et al., (2000). Our replication therefore supports the original findings of a mediating effect from the technical hitch, also in good outcomes. As mentioned earlier, the technical hitch scenarios contradicted Spellman`s probabilistic reasoning theory. The different card scenario rejected the theory that people update probabilistic calculations after each event, as the temporality effect was

eliminated. Arguably, our successful replication of the group differences in mutation therefore provides support for the rejection of the probabilistic reasoning theory. Our addition of responsibility also supported this notion, as participants showed the standard temporality effect in the same card scenario, and an elimination of this tendency in the different card scenario. Thus, our replication findings support the theoretical framework presented by Byrne et al., (2000), that a presupposition of the first event competes with the explicit availabilities of its counterfactuals, making participants equally likely to follow either route when mutating.

However, as mention earlier, our findings on relief were somewhat contradicting to the original findings. Our findings showed that the temporality effect was eliminated in both the scenarios, opposed to only in the different card scenario. Moreover, our calculations of the original effect sizes presents that there was no signal in relief in neither scenario. The replication finds a signal in both scenarios for relief, but the effects are close to no signal (see table 4 for comparison). In general, our findings suggest that there is little to no tendency of a temporality effect for relief in the scenarios. However, there are different possible explanations for these results. One possible explanation is the difference in response options in the studies, as Byrne et al., (2000) included “neither”, while we only included the binary options of the two participants. Looking closer at the original results for attribution of relief, the difference within the same card scenario was reported as somewhat marginal. However, the difference was eliminated within the different card scenario, making the difference between groups notable. When looking at respondents for the “neither” option in the scenarios, there is an 11% increase in the different card, compared with the same card. On one hand one might argue that this increase may contribute to the elimination of a difference between “Jones” and “Brady” in the difference card, thus making the difference between groups larger. On the other hand, the increase also reflects a difference in how the participants

are somewhat more decisive in choosing either Jones or Brady in the same card, perhaps reflecting a more clear cut tendency than in the different card. Either way, the attempt of our replication might be effected to some extent by the exclusion of "neither" as an option. In addition, even though Byrne et al., (2000) found a difference in relief the good outcome experiment, other experiments from the study found a distinction between mutation and emotional attribution. In their discussion, Byrne et al., (2000) argued that emotional attribution might be partly separated from mutability and counterfactuals. Thus, our replication provided somewhat differential findings from the original study. However, they are consistent with the suggestion that the relation between mutability and emotional attribution is not as straightforward as previously assumed (Byrne et al., 2000).

In addition to the replication studies, our extension implementing choice may provide further insight into the temporality effect. In the same card scenario we found no difference in mutation, relief, or responsibility, consistent with an elimination of the hypothesized temporality effect. In addition, we found a reversal of the temporality effect in the different card scenario, across all three questions. These results suggest that choice can have a mediating effect on temporality effect, when giving control to the first person. That choice is of importance can be linked to earlier studies on controllable events, and action versus inaction. Controllable events, decision based actions, have been theorized to be more mutable than uncontrollable events. (Girotto et al., 1991). One explanation for this is because of how a decision is composed of at least two different alternatives, therefore presenting counterfactuals more readily in mind (Girotto et al., 1991). In addition, actions are found to be more mutable than inactions, at least in the short term. Actions are mentally presented with a post- and pre-action state, whereas inactions are presented as just one state. Thus actions explicitly presents more information, making counterfactuals more readily available (Byrne &

McEleney, 2000). Both of these explanations are based on the understanding that availability of counterfactuals increases the mutability of an event.

In the extension scenarios Anne, the first person, is the only one that gets to make an active decision, choosing to either draw again or not. In addition to the decision being an controllable event, she also performs an action as she actively chooses to draw again in both scenarios. Participants might therefore have more counterfactuals for Anne readily in mind, compared with Joan, which makes her event more mutable. Furthermore, these scenarios are based on the technical hitch scenarios in experiment three in Byrne et al., (2000), which, as previously discussed, creates two opposing tendencies of mutability. Namely, that participants base their creation of the counterfactual model on either the presupposition of the first event, or on the represented information from the factual events (Byrne et al., 2000). Notably, the extension is only fitted for a discussive comparison with the original and replication findings, but the same theoretical understandings might be suitable.

Firstly, the elimination in the same marble scenario could, similarly to the different card scenario in the original studies, derive from two opposing tendencies. These tendencies being the presupposition of the first event, opposing the available counterfactuals of the first event. As the original studies did not have available counterfactuals in the same marble scenario, a standard temporality effect occurred. However, counterfactuals were made available here by the alternatives of the decision, arguably acting as the countering tendency against presupposition.

Secondly, the reversal in the different marble scenario may derive from how multiple sources present counterfactuals for the first event, outweighing its presupposition. Arguably, similarly to the different card scenario in the original study, counterfactuals are made explicitly available through factual representation. However, counterfactuals might also here derive from the alternatives of the decision, as well as through the pre- and post-action states.

These multiple sources might influence more participants to create their counterfactual model based on availability, rather than presupposition, perhaps resulting in the reversal rather than the elimination of the standard temporality effect. Thus, with such interpretations of choice implementation, extension findings might be explained with the theoretical framework of the counterfactual model (Walsh & Byrne, 2004).

In contrast, the results are perhaps not as fitting when attempting to use the probabilistic reasoning framework (Spellman, 1997). One contradicting finding is how the technical hitch lead to between group differences. From probabilistic reasoning, participants should update their calculations after each event, and calculate the latest event to have the biggest probabilistic effect on the outcome. However, action versus inaction is mentioned in the theory as a factor that effects causal attributions, with actions being attributed more causality. The reasoning being that performing an action generally changes the probability of the outcome, while inactions does not change the probability (Spellman, 1997). However, Anne`s action does not change the probable chances of winning, as the odds are 50/50 either way. Moreover, availability of alternatives is also brought up as a causality factor, as availability of alternatives before the outcome can have different probabilistic effect on the outcome. However, the alternatives to Anne`s throw does not change the probability of the outcome, as the outcome would still be dependent on Joan`s throw.

Furthermore, the crediting causality hypothesis is generally based on two sequential events. In the technical hitch scenario in the original studies, the technical hitch presented that the show had to restart, implying a discard of the pre-hitch event. In the extension however, the addition of another draw was by choice, making it a third action. As this means an addition of an event, the use of the crediting causality theory is perhaps generally not suited for our extension. However, Segura et al., (2002) found that the temporality and causal order

effects are found in sequences with more than two events as well. In general, the extension might therefore provide some discussive support for the counterfactual model.

Limitations and Future Directions

Regardless of our aim of close replications, our replications had several differences from the original studies. Among these was the difference in sample size, as we increased the sample size in both studies. A larger sample size is however often fruitful, and is especially relevant for replications of earlier studies with small sample sizes. Even so, the sample sizes are still somewhat small, and especially study 2 would benefit with even larger samples. The samples also consisted of a different population, as both the original studies used undergraduate students from English speaking countries. Our studies however consisted of a more varied age group, as well as study 1 consisting of Norwegian speaking participants. Moreover, we made some alterations of the experimental design in study 2, as depicted in the method section of the study. We also did not conduct all statistical analyzes that could be of interest, for instance in relation to relief, and responsibility. Even though we conducted a between group comparison of mutation in study 2, future research could benefit from conducting the test across all questions. Regarding our extension, some limitation is reflected in the binary response options. As making the participants choosing between only two options, we had to eliminate several, plausibly equally relevant, options. In addition, other scenarios could also be included and compared, for instance a scenario where both participants got the choice to redraw or not. Future research would therefore benefit from including more alternatives and scenarios to the marble experiment.

In general, the differences between the original studies and our replications may have some effect on the results. In addition, the extension only yields ground for a conceptual discussion of the temporality effect, excluding the possibility of a statistical comparison. This

limitation could be accounted for in future research, furthering examining choice in temporality effect. Moreover, future research would benefit from conducting replications with larger sample sizes, and further examine the relationship between causal attribution and mutation. In general, our replications provided some additional information for the theoretical discussion of temporality effect. Lastly, part of our findings could be seen as support for the reliability of the original findings, showing some support for the generalizability of the temporality effect.

Conclusion

Overall, our replication findings show a general support of the original study`s findings of a temporality effect. In mutation, guilt, and blame in Miller and Gunasegaram (1990), our replication findings provide support for all three questions. In mutation in Byrne et al., (2000) we also find support for the original findings. Relief was replicated, however the results appeared to show no difference in relief attribution. Moreover, our addition of responsibility showed a temporality effect in line with the original findings of temporality effect. Our extension findings show that giving the first person the possibility to draw again changes, what we would otherwise assume, would be a standard temporality effect scenario. In addition, the findings support the counterfactual model, somewhat more, than the crediting causality model.

References

American Psychological Association. (2020). Publication manual of the American Psychological Association (7th ed.). Washington, DC: Author.

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach,F. J., Geller, J., Giner-Sorolla, R., ... & Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication?.*Journal of Experi-mental Social Psychology*, 50, 217-224. https://doi.org/10.1027/1864-9335/a000178

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, 67, 135-157. https://doi.org/10.1146/annurev-psych-122414-033628

Byrne, R. M. J., & McEleney, A. (2000). Counterfactual Thinking About Actions and Failures to Act. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1318-1331. https://doi.org/10.1037/0278-7393.26.5.1318

Byrne, R. M. J., Segura, S., Culhane, R., Tasso, A., & Berrocal, P. (2000). The temporality effect in counterfactual thinking about what might have been. *Memory & Cognition*, 28(2), 264–281. https://doi.org/10.3758/BF03209313

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge Academic

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, 78(1), 31–48. https://doi.org/10.1016/0001-6918(91)90007-M

Hagger, M. S. (2019). Embracing Open Science and Transparency in Health Psychology. *Health Psychology Review*, 131-136. https://doi.org/10.1080/17437199.2019.1605614

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136. https://doi.org/10.1037/0033-295X.93.2.136

LeBel, E. P., Cheung, I., Vanpaemel, W., & Campbell, L. (2019). A Brief Guide to Evaluate Replications. *Meta-Psychology*, 3, MP.2018.843. https://doi.org/10.15626/MP.2018.843

Miller, D. T., & Gunasegaram, S. (1990). Temporal order and the perceived mutability of events: Implications for blame assignment. *Journal of Personality and Social Psychology*, 59(6), 1111–1118.

Ogloff, J. R. P., & Finkelman, D. (1999). Psychology and Law. I R. Roesch, S. D. Hart & J. R. P. Ogloff (Red.), *Perspectives in Law & Psychology* (Vol. 10, p. 1-20). Springer.

Segura, S., Fernandez-Berrocal, P., & Byrne, R. M. J. (2001). Temporal and causal order effects in thinking about what might have been. *Memory & Cognition*, 29(6), 857–864. https://doi.org/10.3758/BF03200468

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, 69, 487-510. https://doi.org/10.1146/annurev-psych-122216-011845

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General*, 126(4), 323–348 https://doi.org/10.1037/0096-3445.126.4.323

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, 23, 87-102. https://doi.org/10.3758/s13423-015-0892-6

Walsh, J. A., & Byrne, R. M. (2004). Counterfactual thinking: The temporal order effect. *Memory & Cognition*, 32(7), 1264-1279. https://doi.org/10.3758/BF03196842

Walsh, C. R., & Byrne, R. M. J. (2001). A computational model of counterfactual thinking: The temporal order effect. I J. D. Moore & K. Stenning (Red.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (p. 1078-1083). Mahwah, NJ: Lawrence Erlbaum Associates.

Supplementary materials

Replication of Miller and Gunasegaram (1990)

Design facet	Replication study
IV operationalization	Same
DV operationalization	Same
IV stimuli	Same
DV stimuli	Same
Procedural details	Similar
Physical settings	Different
Contextual variables	Different
Replication classification	Close replication

Note: replication summary based on the criteria by LeBel et al., (2019)

Byrne et al., (2000)

Design facet	Replication study
IV operationalization	Same
DV operationalization	Same
IV stimuli	Same
DV stimuli	Same
Procedural details	Similar
Physical settings	Different
Contextual variables	Different
Replication classification	Direct replication

Note: replication summary based on the criteria by LeBel et al., (2019)

Replication Continuum

Highly Similar ← → Highly Dissimilar

	Direct Replication			Conceptual Replication	
Design Facet	Exact Replication (All facets under researcher control are the same)	Very Close Replication (Procedure or physical setting is different)	Close Replication (IV or DV stimuli are different)	Far Replication (IV or DV operationalization or population is different)	Very Far Replication (IV or DV constructs are different)
Effect, Hypothesis	Same	Same	Same	Same	Same
IV Construct	Same	Same	Same	Same	Different
DV Construct	Same	Same	Same	Same	Different
IV Operationalization	Same	Same	Same	Different	
DV Operationalization	Same	Same	Same	Different	
Population (e.g., age)	Same	Same	Same	Different	
IV Stimuli	Same	Same	Different		
DV Stimuli	Same	Same	Different		
Procedural Details	Same	Different			
Physical Setting	Same	Different			
Contextual Variables	Different				
⋮	⋮				

Fig. 1. Taxonomy for classifying a replication study's methodological similarity to an original study. "Same" indicates that the design facet in question is the same as in the original study, and "different" indicates that it is different. IV = independent variable; DV = dependent variable. "Population" refers to major population characteristics, such as age and whether the sample is drawn from the community or a special clinical population. Procedural details are minor experimental particulars (e.g., task instructions, font, font size). Contextual variables are design facets beyond a researcher's control (e.g., history, culture, language).

Note: Table for replication criteria by LeBel et al., (2019)

Sikt

Assessment of processing of personal data

Reference number
754105

Assessment type
Standard

Date
08.03.2023

Project title
Temporality effect and counterfactual thinking

Data controller (institution responsible for the project)
Norges teknisk-naturvitenskapelige universitet / Fakultet for samfunns- og utdanningsvitenskap (SU) / Institutt for psykologi

Project leader
SUBRAMANYA PRASAD CHANDRASHEKAR

Project period
15.01.2023 - 30.06.2023

Categories of personal data
General

Legal basis
Consent (General Data Protection Regulation art. 6 nr. 1 a)

The processing of personal data is lawful, so long as it is carried out as stated in the notification form. The legal basis is valid until 30.06.2023.

Notification Form

Comment
Data Protection Services has assessed the changes registered on 02.03.2023.

Another sample has been added to the notification form (sample 2). Processing is based on informed consent.

We find that the processing of personal data in this project is lawful and complies with data protection legislation, so long as it is carried out as described in the Notification Form with dialogue and attachments.

FOLLOW-UP OF THE PROJECT
We will follow up the project at the planned end date in order to determine whether the processing of personal data has been concluded.

Good luck with the project!

Project working title: **Conceptual replication of good outcomes experiment 3 in Byrne et al. (2000): The temporality effect in counterfactual thinking about what might have been.**

Authors: ***left out for blind review of the preregistration***
Affiliation: ***left out for blind review of the preregistration***

Summary

This project's aim is to test whether there is a temporality effect, the predominant tendency people have to consider the second event in an independent two-event sequence more mutable than the first, for situations with good outcomes.

Hypotheses

Predictions across both Scenarios (Same and different card scenarios)
H1: The most recent event in a good outcome sequence of two independent events is considered more mutable than the first.

H2: Given a positive outcome based on a sequence of two independent chance events, participants will judge the second player to experience more relief than the first player.

H3: Given a positive outcome based on a sequence of two independent chance events, participants will judge the second player to feel more responsible for winning.

We will test these hypotheses across both same and different card conditions.

Study Materials

Study Outline:

In this survey, you will read a scenario concerning two participants drawing cards. Following the scenario, you will be asked to answer some simple questions, please answer them in the order they are asked, and refrain from changing your answer after writing it down.

Scenario 1- same card:
Imagine two individuals (Jones and Brady) who take part in a television game show, on which they are offered the following very attractive proposition. Each individual is given a shuffled deck of cards, and each one picks a card from his own deck. If the two cards they pick are of the same color (i.e., both from black suits or both from red suits), each individual wins £1,000. However, if the two cards are not the same color, neither individual wins anything.

Jones goes first and picks a red card from his deck. At this point, the game show host has to stop the game because of a technical difficulty. After a few minutes, the technical problem is

solved and the game can be restarted. Jones goes first again, and this time he draws a red card. Brady goes next and the card that he draws is also a red card. Thus, the outcome is that both individuals wins £1,000 each.

Question 1: After the draw both Jones and Brady reflected on how lucky they had been. After all, if one of them had picked a different card they might neither have won the £ I,000; for instance, if...:

Please choose one of the following options:

- Jones drew a black card the second time
- Brady drew a black card

Question 2: Who would you predict would experience more relief at having won?

Please choose one of the following options:

- Jones
- Brady

Question 3: Who would you predict would feel more responsible for the winning outcome?

Please choose one of the following options:

- Jones
- Brady

Scenario 2- different card:

Imagine two individuals (Jones and Brady) who take part in a television game show, on which they are offered the following very attractive proposition. Each individual is given a shuffled deck of cards, and each one picks a card from his own deck. If the two cards they pick are of the same color (i.e., both from black suits or both from red suits), each individual wins £1,000. However, if the two cards are not the same color, neither individual wins anything.

Jones goes first and picks a black card from his deck. At this point, the game show host has to stop the game because of a technical difficulty. After a few minutes, the technical problem is solved and the game can be restarted. Jones goes first again, and this time the card that he draws is a red card. Brady goes next and the card that he draws is also a red card. Thus, the outcome is that both individuals wins £1,000 each.

Question 1: After the draw both Jones and Brady reflected on how lucky they had been. After all, if one of them had picked a different card they might neither have won the £ I,000; for instance, if...:

Please choose one of the following options:

- Jones drew a black card the second time
- Brady drew a black card

Question 2: Who would you predict would experience more relief at having won?

Please choose one of the following options:

- Jones
- Brady

Question 3: Who would you predict would feel more responsible for the winning outcome?

Please choose one of the following options:

- Jones
- Brady

Demographic questions

Thank you, you completed the main part of the survey.

A couple of quick final questions.

How old are you?

---- (text box here)

What is your gender?

- Female
- Male
- Other

Planned sample

Participants recruited will be English-speaking adults from Prolific. The sample size was calculated using G*power 3.1.9.7, and was based on 90% power (and $\alpha = .05$), with the aim of detecting an effect size of Cohen`s $g = -0.012$. We aim to achieve a sample size of 300 particpants. The details from the power analysis can be found in the power analysis section below.

Suggested Analysis

The original authors used a one sample binomial test, and the replication will use the same test as it is suited for the respective hypotheses.

Detailed results of card scenario by Byrne et al. (2000)

Table 1
Percentages and frequencies measures

Question	In %	Conversion	After rounding
Same card			
Undoing choice			
First only	17%	7,82	8
First then second	9%	4,14	4
Overall first	26%	11,96	12
Second only	50%	23	23
Second then first	9%	4,14	4
Overall second	59%	27,14	27
Total	85%	39,1	39
Who will feel more relief			
First	35	16,1	16
Second	59	27,14	27
Neither	2	0.92	1
n	46		

Different card			
Undoing choice			
First only	33	15,84	16
First then second	4	1.92	2
Overall first	37.5	18	18
Second only	31	14.88	15
Second then first	2	0.96	1
Overall second	33	15.84	16
Total	70.5	33.84	34
Who will feel more relief?			
First	44	21.12	21
Second	42	20.16	20
Neither	13	6.24	6
n	48		

Note: Frequency calculations were based on percentage values reported in the original study (Byrne et al., 2000, s. 270)

Table 2

Results from original study by Byrne et al. 2000

Hypotheses	**Dependent variables**	**Statistical test**	**Effect size with 95% CI**
	Same-card	**One sample Binomial-test**	**Cohen`s *g***
H1	Second rather than first mutable	59% vs. 26%, binomial *n* = 39, k= 12, *p* < .01	0.192 [0.041, 0.344]
H2	Second rather than first relief	59% vs. 35%, binomial *n* = 43, k=16, *p* < .06	0.128 [-0.023, 0.278]
	Different-card		
H1	No temporality effect mutable	33% vs. 37.5%, binomial *n* = 34, k = 16, *p* < .43	-0.029 [-0.206, 0.147]
H2	No temporality effect relief	42% vs. 44%, binomial *n* = 41, k = 20, z = 0, *p* < .50	-0.012 [-0.172, 0.148]

Note: These calculations are based on overall proportions.

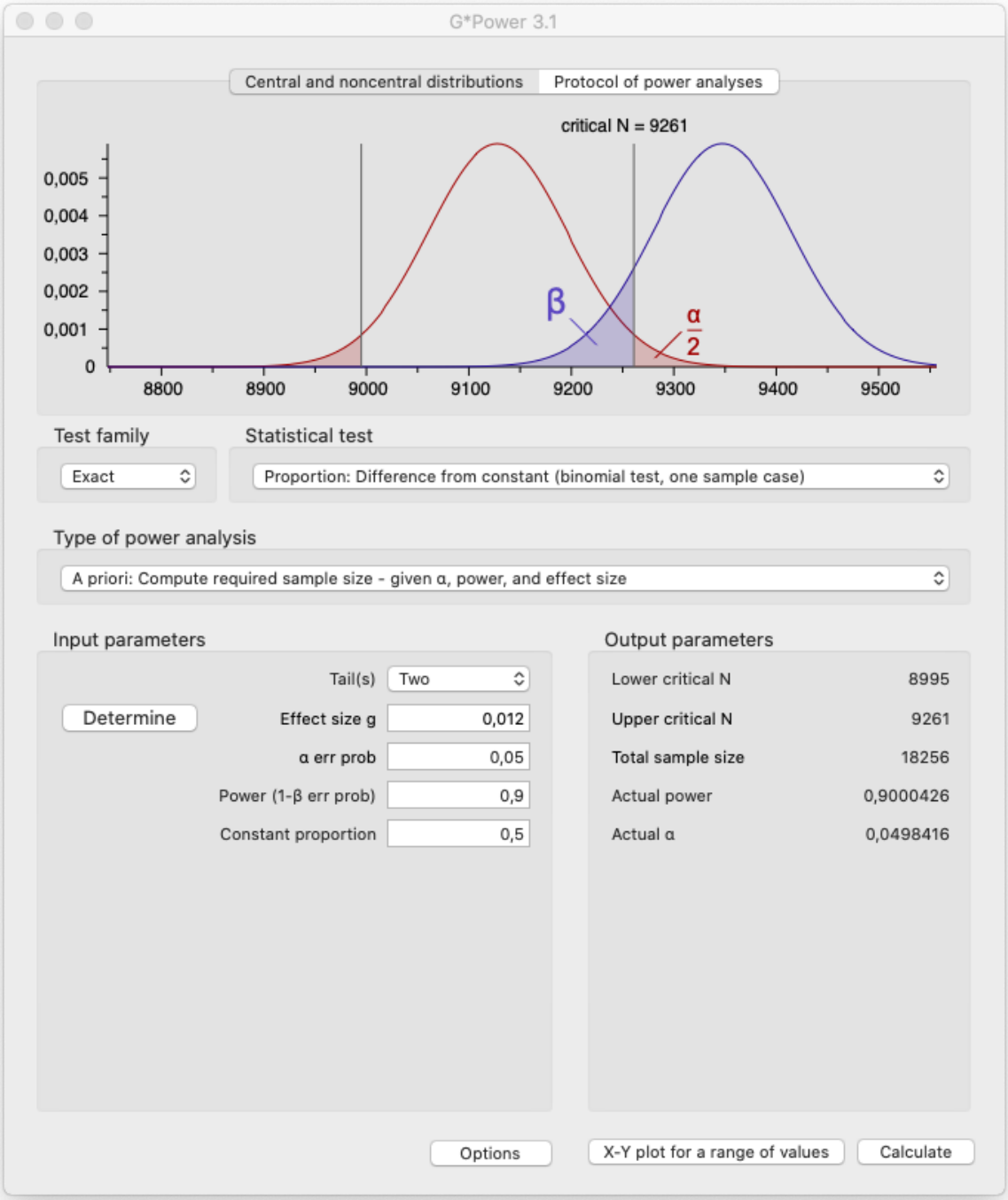
Power Analysis

The rationale for reconstructing the original dataset and re-running analysis: authors of the original studies did not report the full statistical results necessary (i.e., effect size measures were missing) to run a power analysis. Hence, we had to re-conduct the analysis reported in the original study based on information available in the description of the study (see Table 1 above).

Steps for power analysis

Hypotheses H1 and H2

The testing of H1 and H2 involved three One sample Binomial test. We conducted a power analysis based on the smallest effect size among these results (i.e., Cohen's g = -0.012). The results of the analysis suggest a total sample size of 18256 (see screenshot below).



Summary of power analysis:
The analysis suggests a replication sample size of 18256 participants. However, we instead aim to recruit a minimum of 300 participants, because of limited resources making it hard to meet the suggested sample size.

Project working title: **Extension of temporality effect**

Authors: ***left out for blind review of the preregistration***
Affiliation: ***left out for blind review of the preregistration***

Summary

This project's aim is to test whether there is a temporality effect, the predominant tendency people have to consider the second event in an independent two-event sequence more mutable than the first, for situations with good outcomes implementing choice.

Hypotheses

Predictions across both scenarios (Same and different marble scenario)

H1: The most recent event in a good outcome sequence of two independent events is considered more mutable than the first.

H2: Given a positive outcome based on a sequence of two independent chance events including choice, participants will judge the second player to experience more relief than the first player.

H3: Given a positive outcome based on a sequence of two independent chance events including choice, participants will judge the second player to feel more responsible for winning.

We will test these hypotheses across both same and different marble conditions.

Study Materials

Study Outline:

In this survey, you will read a scenario concerning two participants drawing marbles. Following the scenario, you will be asked to answer some simple questions, please

answer them in the order they are asked, and refrain from changing your answer after writing it down.

Scenario 1- same marble:

Imagine two individuals, Anne and Joan, who are offered the following proposition. Each individual is given a sack of marbles, and each one draws a marble from her own sack. If the two marbles they draw are of the same color (i.e., both are blue or both are white), each individual wins £1,000. However, if the two marbles are not the same color, neither individual wins anything.

Anne goes first and picks a blue marble from her sack. At this point Anne gets the option to choose to either draw again, or to not draw again and keep her blue marble. She chooses to draw again and draws a blue marble. Joan goes next and the marble she draws is a blue marble. Thus, the outcome is that both individuals wins £1,000 each.

Question 1: After the draw both Anne and Joan reflected on how lucky they had been. After all, if one of them had picked a different marble they might neither have won the £ I,000; for instance, if...:

Please choose one of the following options:

- Anne drew a white marble the second time
- Joan drew a white marble

Question 2: Who would you predict would experience more relief at having won?
Please choose one of the following options:

- Anne
- Joan

Question 3: Who would you predict would feel more responsible for the winning outcome?

Please choose one of the following options:

- Anne
- Joan

Scenario 2- different marble:

Imagine two individuals, Anne and Joan, who are offered the following proposition. Each individual is given a sack of marbles, and each one draws a marble from her own sack. If the two marbles they draw are of the same color (i.e., both are blue or both are white), each individual wins £1,000. However, if the two marbles are not the same color, neither individual wins anything.

Anne goes first and picks a white marble from her sack. At this point Anne gets the option to choose to either draw again, or to not draw again and keep her white marble. She chooses to draw again and this times she draws a blue marble. Joan goes next and the marble she draws is a blue marble. Thus, the outcome is that both individuals wins £1,000 each.

Question 1: After the draw both Anne and Joan reflected on how lucky they had been. After all, if one of them had picked a different marble they might neither have won the £ I,000; for instance, if...:

Please choose one of the following options:

- Anne chose to not draw again
- Joan drew a white marble

Question 2: Who would you predict would experience more relief at having won?
Please choose one of the following options:

- Anne

- Joan

Question 3: Who would you predict would feel more responsible for the winning outcome?

Please choose one of the following options:

- Anne
- Joan

Demographic questions

Thank you, you completed the main part of the survey.

A couple of quick final questions.

How old are you?

---- (text box here)

What is your gender?

- Female
- Male
- Other

Planned sample

Participants recruited will be English-speaking adults from Prolific. The sample size was calculated using G*power 3.1.9.7, and was based on 90% power (and $\alpha = .05$), with the aim of detecting an effect size of Cohen`s $g = 0.1$. We aim to achieve a

sample size of 300 participants. The details from the power analysis can be found in the power analysis section below.

Suggested Analysis

Hypotheses H1, H2, H3

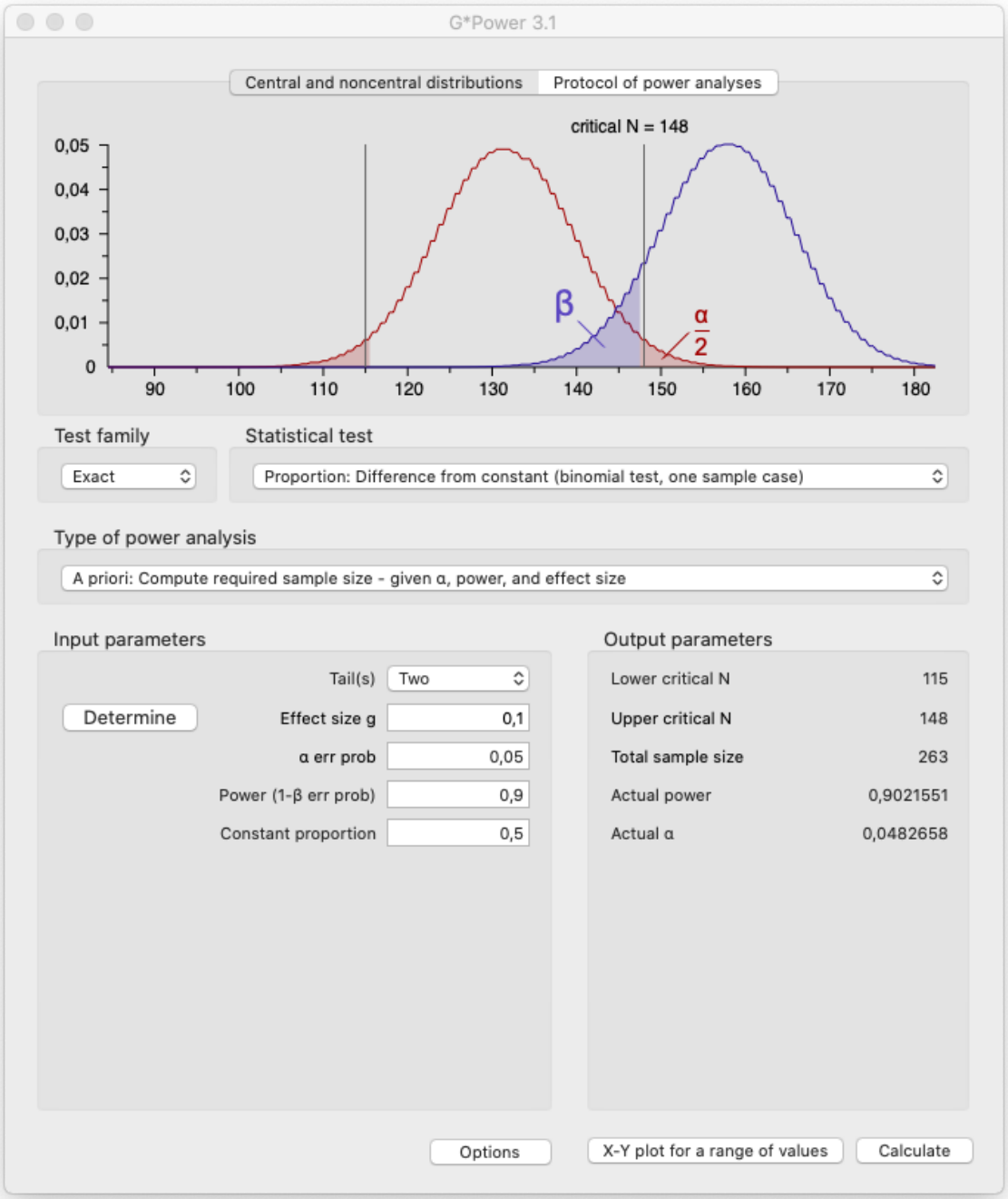
The testing of H1, H2 and H3 will involve three One sample Binomial tests.

Power Analysis

Steps and summary for power analysis

Smallest effect size of interest

We aim to be able to detect a small effect size of Cohen`s g = 0.1 at 0.90 power (alpha = .05). The result of the power analysis suggests a total sample size of 263 (see screenshot below), but we will aim for 300 participants.


G*Power 3.1
Central and noncentral distributions
Protocol of power analyses
critical N = 148
0,05
0,04
0,03
0,02
0,01
0
90
100
110
120
130
140
150
160
170
180
β
α/2
Test family
Exact
Statistical test
Proportion: Difference from constant (binomial test, one sample case)
Type of power analysis
A priori: Compute required sample size - given α, power, and effect size
Input parameters
Tail(s)
Two
Determine
Effect size g
0,1
α err prob
0,05
Power (1-β err prob)
0,9
Constant proportion
0,5
Output parameters
Lower critical N
115
Upper critical N
148
Total sample size
263
Actual power
0,9021551
Actual α
0,0482658
Options
X-Y plot for a range of values
Calculate

Project working title: **Replication of Miller & Gunasegaram`s (1990) coin toss scenario**

Authors: ***left out for blind review of the preregistration***
Affiliation: ***left out for blind review of the preregistration***

Summary

This project's aim is to test the predominant tendency people have to consider the second event in an independent two-event sequence more mutable than the first, and to test whether linguistic preferences influence this tendency in a Norwegian sample.

Hypotheses

Common predictions across Scenario 1 and Scenario 2

H1: The most recent event in a sequence of two independent events is considered more mutable than the first.

H2: Given a negative outcome based on a sequence of two independent chance events, participants will judge the second player to experience more guilt than the first player.

H3: Given a negative outcome based on a sequence of two independent chance events, the study participants will judge that the first player will blame the second player more often than the second player blaming the first player.

Contrasting Scenario 1 and Scenario 2

We do not have concrete directional predictions contrasting the responses across Scenario 1 and Scenario 2. We will report the results of the contrast in responses to Scenario 1 and Scenario 2 across each of the three predictions noted above (i.e., H1, H2, & H3).

Exploratory predictions:

The survey will be conducted in the Norwegian language. We will test if the Norwegian participants report “heads-tails” linguistic preference over “tails-heads,” and if such a preference contributes to the pattern of responses across Scenario 1 and Scenario 2.

Study Materials

Study Outline:

In this survey, you will read two separate scenarios about outcomes of chance events (such as a coin toss). As you read the descriptions, please carefully try to form a detailed understanding of the situations related to the decision-makers involved. Following each scenario, you will answer three short questions about the scenario based on your understanding.

Scenario 1:

Imagine two individuals (Jonas and Kristian) who are offered the following very attractive proposition. Each individual is asked to toss a coin. If the two coins come up the same (both heads or both tails), each individual wins NOK 10,000. However, if the two coins do not come up the same, neither individual wins anything. Jonas goes first and tosses a head; Kristian goes next and tosses a tail. Thus, the outcome is that neither individual wins anything.

There were two ways that Jonas & Kristian could have won NOK 10,000. Which of these alternatives comes more readily to mind?

- Jonas tossing a tail
- Kristian tossing a head

Who would you predict will experience more guilt—Jonas or Kristian?

- Jonas
- Kristian

Will Jonas blame Kristian more or will Kristian blame Jonas more for their failure to win NOK 10,000?

- Jonas
- Kristian

Scenario 2:

Imagine two individuals (Oscar and Chris) who are offered the following very attractive proposition. Each individual is asked to toss a coin. If the two coins come up the same (both heads or both tails), each individual wins NOK 10,000. However, if the two coins do not come up the same, neither individual win anything. Oscar goes first and tosses a tail; Chris goes next and tosses a head. Thus, the outcome is that neither individual win anything.

There were two ways that Oscar & Chris could have won NOK 10,000. Which of these alternatives comes more readily to mind?

- Oscar tossing a head
- Chris tossing a tail

Who would you predict will experience more guilt—Oscar or Chris?

- Oscar
- Chris

Will Oscar blame Chris more or will Chris blame Oscar more for their failure to win NOK 10,000?

- Oscar will blame Chris more
- Chris blame Oscar more

Which of these phrasings appear the most natural to you?

- Heads or tails
- Tails or heads

Demographic questions

Thank you, you completed the main part of the survey.

A couple of quick final questions.

How old are you?

---- (text box here)

What is your gender?

- Female
- Male
- Other

Planned sample

Participants recruited will be Norwegian-speaking adults. The sample size was calculated using G*power 3.1.9.7, and was based on 90% power (and α = .05), with the aim of detecting an effect size of Cohen`s *g* = 0.2. We aim to achieve a sample size of 200 particpants. The details from the power analysis can be found in the power analysis section below.

Suggested Analysis

The original authors did not include the results of the analysis, nor did they report which one proportion test they used, however, we found it appropriate to conduct one sample Z tests to test the predictions.

Detailed results of coin toss scenario by Miller & Gunasegaram (1990)

Table 1

Percentages and frequency count measures on predictions of experience of guilt, judgment of blame, and undoings of the first or second sequence through an "if only…" question.

Question	In %	Conversion	After rounding
Who will experience more guilt?			
First	14%	12,32	12
Second	86%	75,68	76
Who will blame the other more?			
First	92%	80,96	81
Second	8%	7,04	7
Which event is most often mutated?			
First event	11%	9,68	10
Second event	89%	78,32	78
n		88	

Note. Frequency calculations were based on percentage values reported in the original study (Miller & Gunasegaram, 1990, pp. 1111-1112).

Table 2

Results of Study 1 of Miller & Gunasegaram (1990)

Hypotheses	Dependent variables	Statistical test	Effect size with 95% CI
		One sample Z-test	**Cohen`s *g***
H1	**DV1:** Which event is most often mutated?	$k = 78$, $n = 88$, $z = 7.25$, $p < .001$	0.39 [0.32 , 0.45]
H2	**DV2:** Who will experience more guilt?	$k = 76$, $n = 88$, $z = 6.82$, $p < .001$	**0.36 [0.29 , 0.44]**
H3	**DV3**: Who will blame the other more?	$k = 81$, $n = 88$, $z = 7.89$, $p < .001$	0.42 [0.36 , 0.48]

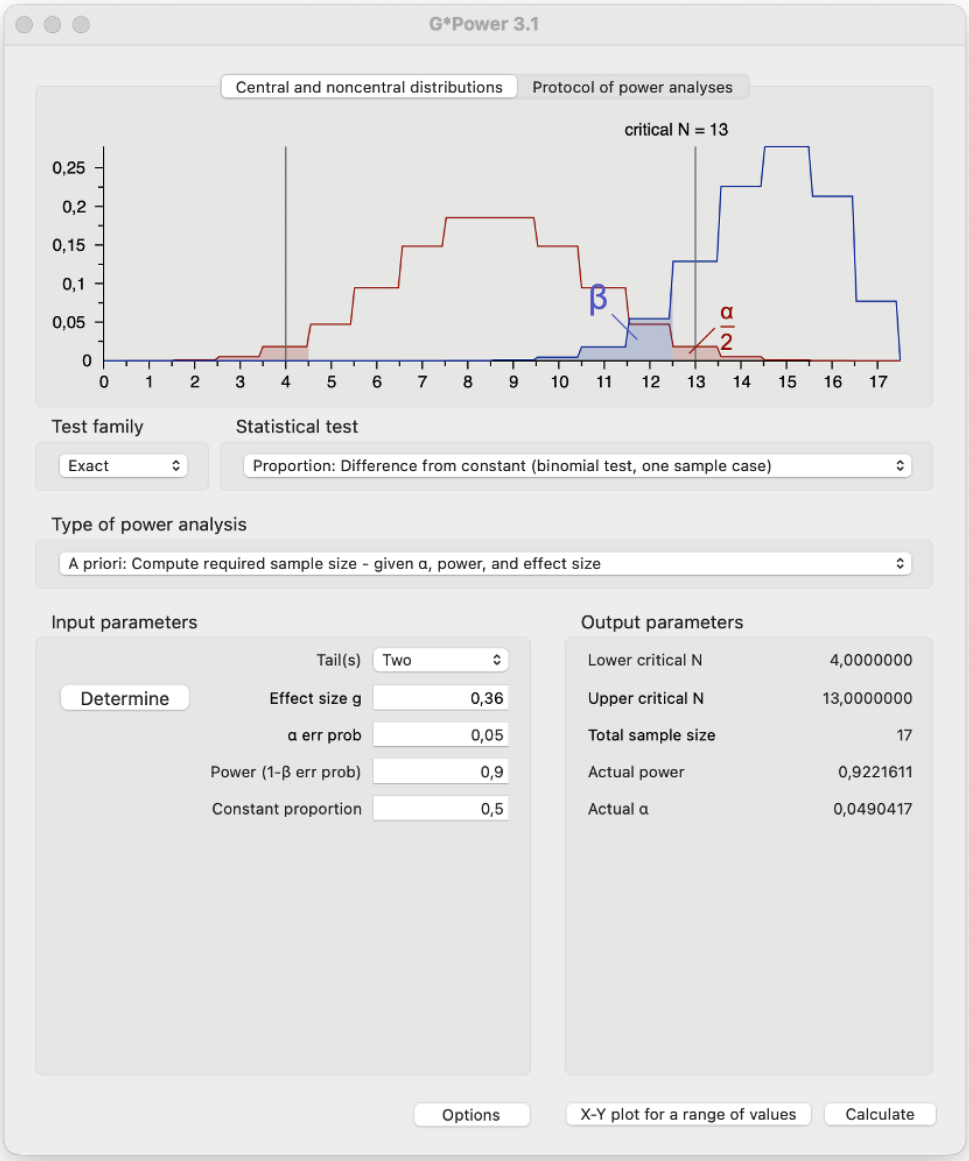
Power analysis

The rationale for reconstructing the original dataset and re-running analysis: authors of the original studies did not report the full statistical results necessary (i.e., effect size measures were missing) to run a power analysis. Hence, we had to re-conduct the analysis reported in the original study based on information available in the description of the study (see Table 1 above).

Steps for power analysis

Hypotheses H1, H2 and H3

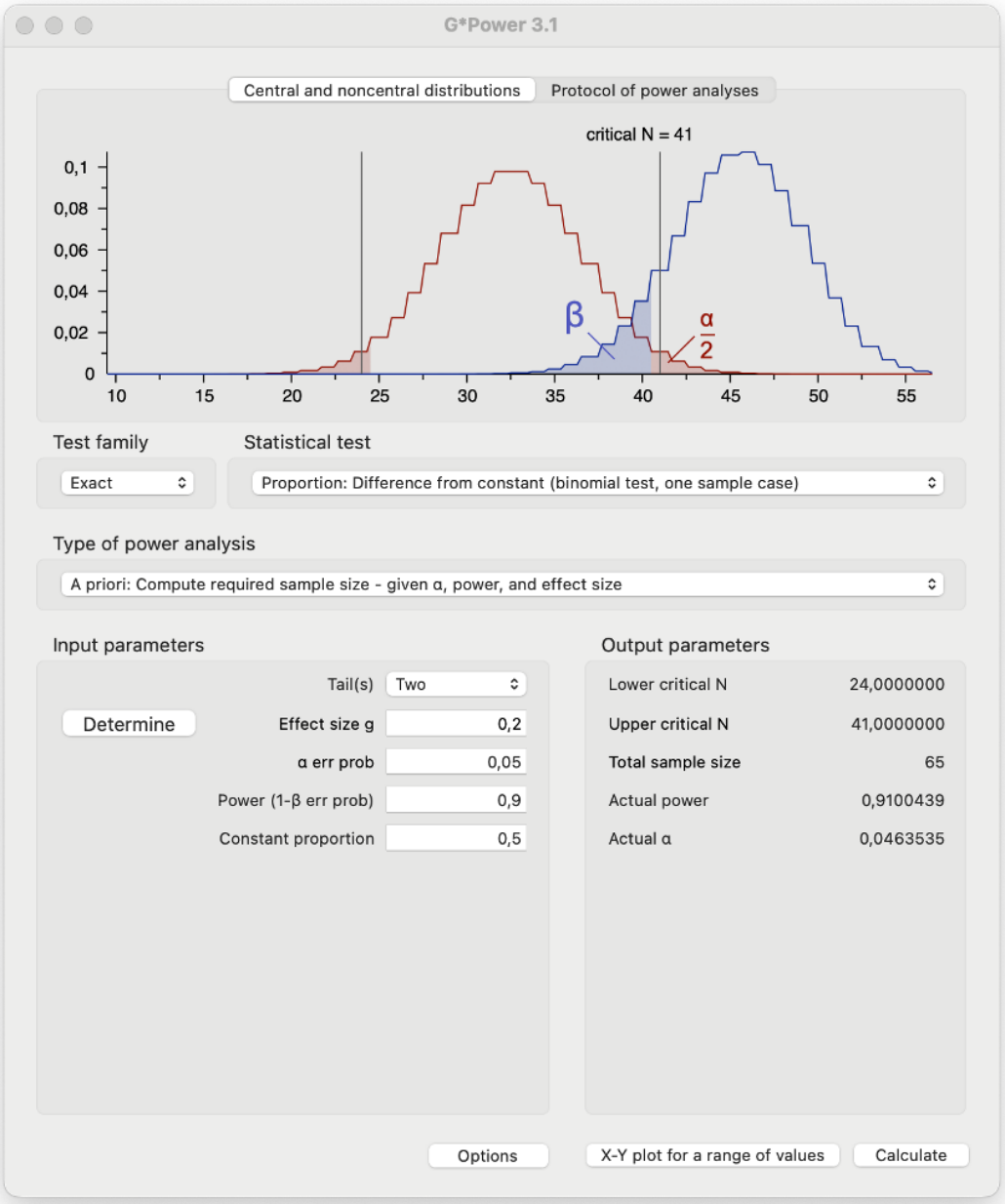
The testing of H1, H2 and H3 involved three One sample Z test. We conducted a power analysis based on the smallest effect size among these results (i.e., Cohen's g = 0.36). The results of the analysis suggest a total sample size of 17 (see screenshot below).



Smallest effect size of interest

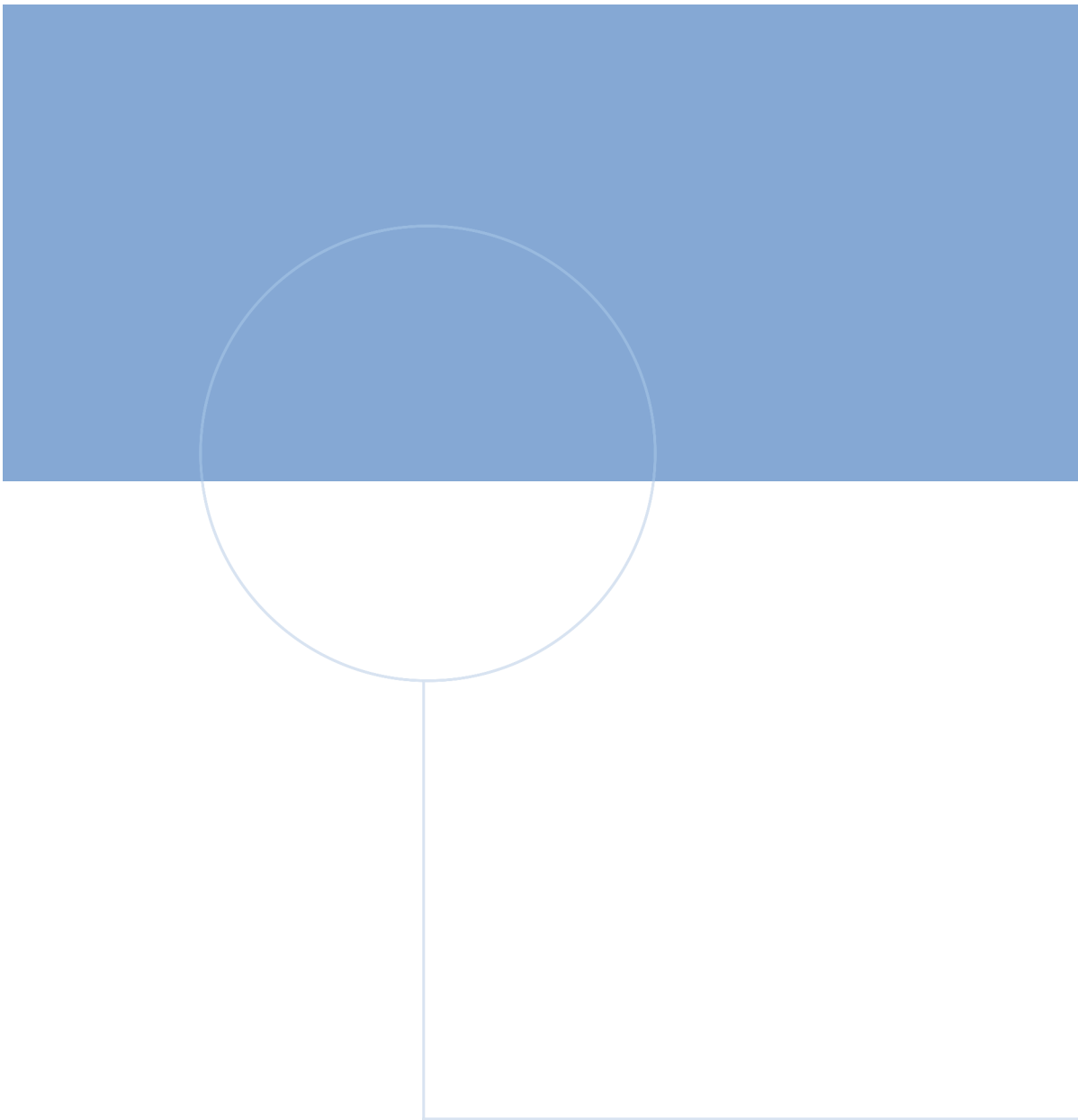
We aim to be able to detect a small effect size of Cohen`s g = 0.2 at 0.90 power (alpha =.05). The result of the power analysis suggests a total sample size of 65 (see screenshot

below).



Summary of power analysis:

The analysis suggests a replication sample size of 65 participants. However, we aim to recruit a minimum of 200 participants, because a larger sample size will be more accurate and reliable. The original study also had very large effect sizes, so to include a larger sample will be beneficial to detect smaller effect sizes.



NTNU

Norwegian University of
Science and Technology