Revisiting Partition Priming in judgment under uncertainty:   
Replication and extension Registered Report of Fox and Rottenstreich (2003)  
 [Stage 1]

Kerou Ding  
ORCID: 0009-0006-9295-0215  
Department of Psychology, University of Hong Kong, Hong Kong SAR  
[u3624657@connect.hku.hk](mailto:u3624657@connect.hku.hk) ; [keruoding@qq.com](mailto:keruoding@qq.com)

^Gilad Feldman  
OCRID: 0000-0003-2812-6599  
Department of Psychology, University of Hong Kong, Hong Kong SAR  
[gfeldman@hku.hk](mailto:gfeldman@hku.hk) ; [giladfel@gmail.com](mailto:giladfel@gmail.com)

^Corresponding author

## Author bios:

Kerou Ding was a thesis student at the University of Hong Kong during 2023-4.

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on judgment and decision-making.

## Declaration of Conflict of Interest:

The author(s) declared no potential conflicts of interests with respect to the authorship and/orpublication of this article.

## Financial disclosure/funding:

The project is supported by the University of Hong Kong Teaching Development Grant.

## Authorship declaration:

Kerou Ding conducted the project as part of her thesis in psychology.

Gilad Feldman guided the project, supervised each step in the project, ran data collection, and edited the manuscript for submission.

## Corresponding author

Gilad Feldman, Department of Psychology, University of Hong Kong, Hong Kong SAR; [gfeldman@hku.hk](mailto:gfeldman@hku.hk) ; 0000-0003-2812-6599

## Rights:

CC BY or equivalent license is applied to the AAM arising from this submission. ([clarification](https://bit.ly/rrs-primer))

## Citation of the target research article:

Fox, C. R., & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science*, *14*(3), 195-200.

## Acknowledgements (with ORCID)

## Contributor Roles Taxonomy

|  |  |  |
| --- | --- | --- |
| **Role** | **Kerou Ding** | **Gilad Feldman** |
| Conceptualization | X | X |
| Pre-registration | X | X |
| Data curation |  | X |
| Formal analysis | X |  |
| Funding acquisition |  | X |
| Investigation | X |  |
| Pre-registration peer review / verification |  | X |
| Data analysis peer review / verification |  | X |
| Methodology | X |  |
| Project administration |  | X |
| Resources |  | X |
| Software | X |  |
| Supervision |  | X |
| Validation |  | X |
| Visualization | X |  |
| Writing-original draft | X |  |
| Writing-review and editing |  | X |

# PCIRR-Study Design Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Statistical tests rationale** | **Interpretation given outcomes** | **Theory affected by the outcomes** |
| Studies 1a and 1b: Do people utilize partition-dependent ignorance priors when forming probability judgments under ignorance or near ignorance? | Studies 1a and 1b: Individuals  utilize partition-dependent ignorance priors when forming probability judgments under ignorance or near ignorance. | We aimed to recruit 600 US Americans through Prolific. See details in the Method’s power and sensitivity analyses subsection.  We calculated Cohen’s h/w and Cohen’s d using the data provided in the target article. | Mann-Whitney U test  Welch’s t-test  Chi squared test | This is a replication.  We mostly followed the statistical analyses of the target article, with some improvements and adjustments. | We will evaluate the replicability of our findings against the target article’s finding using the LeBel et al. (2019) paradigm (examining signal and comparison of confidence intervals with the target article’s effect size). | The principle of insufficient reason.  Partition dependence. |
| Study 3: Do people utilize partition-dependent ignorance priors when forming probability judgments under uncertainty? | Study 3: Individuals utilize partition-dependent probabilities when forming judgments under uncertainty (1/2 > 1/n). | Mann-Whitney U test  Welch’s t-test  Chi squared test |
| Study 4: Do people exhibit a tendency to rely on partition-dependent ignorance priors when making decisions? | Study 4: Individuals utilize partition-dependent probabilities when making decisions under ignorance (1/2 > 1/n). | Two proportion z-test (/chi-square) |

# Abstract

[IMPORTANT: Abstract, method, and results were written using a randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection. For the purpose of the simulation, we wrote things in past tense, but no pre-registration or data collection took place yet.]

Partition dependence is the phenomenon in which individuals' evaluations of probabilities are influenced by the partitioning of the information, based on how the information is presented or framed. In a Registered Report experiment involving an American online Amazon Mechanical Turk sample (*N* = 600), we conducted a replication and extension of Studies 1a, 1b, 3, and 4 from a classic article by Fox and Rottenstreich (2003) demonstrating the phenomenon. They showed that participants’ estimations of an event’s likelihood shifted based on minor adjustments of the framing that suggested a change in partitioning perspective (“ignorance prior” priming). [The following is a demo placeholder and will be updated following data collection.] We [found/did not find] support for Fox and Rottenstreich (2003)’s findings [...] Materials, data, and code are available on: <https://osf.io/g9czs/>

*Keywords*: Bias, judgment and decision making, partition dependence, ignorance prior, risk, registered replication.

# 

# Revisiting Partition Priming in judgment under uncertainty: Replication and extension Registered Report of Fox and Rottenstreich (2003) [Stage 1]

[IMPORTANT: Section is written in the past tense to simulate what the manuscript will look like after data collection, yet no pre-registration or data collection took place yet.]

## Background

When lacking specific knowledge about a situation, people tend to rely on an “ignorance prior”, an assumption that assigns equal chances to different possibilities. Yet, the perceived ignorance prior may depend on the way the information is presented, and how the person mentally maps the available possibilities. Fox and Rottenstreich (2003) demonstrated that a slight change in the framing of the same problem, partitioning, may lead to the use of different ignorance priors and therefore shift evaluations of perceptions of the probabilities. For example, when asked about the weather the following week with no relevant information, the questions “will Sunday be the hottest day next week?” and "will Sunday be hotter than all other days next week? " might trigger different ignorance priors. The first might trigger a 1/2 versus 1/2 given a mental split of “hotter than rest of the week” vs. “not hotter”, whereas the latter might trigger a 1/7 vs. 6/7 given a mental comparison of Sunday against each of the other six days.

We conducted an independent close replication and extension Registered Report of Fox and Rottenstreich (2003), as one of the classic articles of the partition dependence phenomenon. We begin by introducing the literature on ignorance priors and the chosen article for replication - Fox and Rottenstreich (2003). We then discuss our motivations for the current replication, review Fox and Rottenstreich (2003) and outline the chosen studies from the target article, their experimental design, and our adaptations and extensions.

## Ignorance prior and partition dependence

Decision making and risk analysis often involve the assessment of probabilities for uncertain events. Individuals rely on the perceived number of potential events in making likelihood estimations when no other information is provided (Johnson-Laird et al., 1999; Fox & Rottenstreich, 2003; Fox & Clemen, 2005). For example, if there are 10 candidates in a competition and no information is available regarding their abilities, one might assign a probability of 1/10 to each candidate's chance of winning the competition. In this case, 1/10 is referred to as the “ignorance prior” of each candidate winning the competition. An ignorance prior denotes a default judgment that assigns equal probabilities to all potential possibilities in a given query, in the absence of any other relevant information.

While the reliance on ignorance priors may initially seem intuitively appealing, it may be sensitive to presentation and framing (Fox & Rottenstreich, 2003; Fox & Clemen, 2005). To illustrate, consider the competition example previously discussed. Imagine an assessor, lacking detailed information on the competition's difficulty and scoring system, is tasked with judging the likelihood that the competition winner receives a score lower than 50. In this scenario, the question tends to elicit a binary partition: {score lower than 50; score higher than 50}. However, if the task is framed so that the task is to estimate the likelihood that the winner's score falls between 50 and 100, then the partition tends to elicit a three-fold partition: {score lower than 50; score between 50 and 100; score higher than 100}. Consequently, this partitioning increases the likelihood that each of those three possibilities is assigned an ignorance prior of 1/3. This adjustment in how the state space is partitioned might shift the assessed probability (Fox & Rottenstreich, 2003; Fox & Clemen, 2005; See et al., 2006). In another word, the default ignorance prior assigned by an assessor is partition-dependent.

The influence of partition dependence on decision making has been observed across different contexts. Fox et al. (2005) reported that the subjective grouping of options influences decision making and resource allocation, recently successfully replicated by Li and Feldman (2024). For example, in a study where participants were asked to select candies from three bowls containing a total of four types of candy, they tended to spread their choices not only across the bowls but also across the different types of candy. This suggested that the distribution of choices was influenced by the way the candies were subjectively grouped or partitioned into different bowls by the experimenters. Feng et al. (2020) showed that when job applicants were partitioned on the basis of various criteria such as gender, nationality, and university, individuals tended to select a more diverse pool of candidates within each specific partitioned criterion. This effect was observed even among experts (seasoned human resource professionals).

According to Fox and Rottenstreich (2003), while the assumption of the principle of insufficient reason is the absence of any relevant information favoring one outcome over another, individuals often possess some relevant knowledge in many real-life situations. In such cases, people may resort to a combination of evaluative strategies and the principle of insufficient reason. Thus, individuals' judged probabilities may exhibit a consistent bias towards partition dependence under conditions of both ignorance (complete absence of relevant knowledge) and uncertainty (possession of relevant knowledge to some extent).

## Choice of study for replication: Fox and Rottenstreich (2003)

We embarked on a replication and extension Registered Report of Fox and Rottenstreich (2003). We aimed to revisit the phenomenon to examine the reproducibility and replicability of the findings with an independent pre-registered well-powered replication and extension. This follows the recent growing recognition of the importance of reproducibility and replicability in psychological science (e.g., Nosek et al., 2022; Zwaan et al., 2018).

We chose Fox and Rottenstreich (2003) based on several factors: its academic and practical impact, potential for improvements in methodology, and mixed findings in the literature.

Fox and Rottenstreich (2003)’s findings have had an impact on scholarly research in the domains of judgment bias, subjective probability estimation, and risk assessment. At the time of writing (March, 2024), there were 232 Google Scholar citations of the article and many important follow-up theoretical and empirical articles, such as Fox and Clemen (2005) and See et al. (2006)’s research on partition-dependent ignorance priors, exploring how these priors were influenced by individuals’ expertise and confidence levels.

There is some indication for mixed findings in the partition dependence literature. For example, Reichelson et al. (2017) replicated the candy bowl study previously conducted by Fox et al. (2005), recruiting both children and adults. They concluded no evidence supporting the impact of candy partition on participants' choices, inconsistent with the findings of Fox et al. (2005). Their findings are in contrast to our successful replication of the same article as a Registered Report (Li & Feldman, 2024).

We attempted to analyze Fox and Rottenstreich (2003)’s studies. The methods and the statistical tests varied across the different studies, despite having very similar designs and hypotheses, with inconsistent reporting of key details for both the procedures and the results. Their brief reporting proved a challenge in deducing the procedure and materials for a replication and comparing the different scenarios. We believe that revisiting classics like our target with reproductions and replications helps clarify important needed details and increase the quality of the literature building on these findings.

## Selection of studies to replicate: Studies 1a, 1b, 3, and 4

Fox and Rottenstreich (2003) investigated ignorance priors and partition dependence in five studies (Studies 1a, 1b, 2, 3, and 4). Studies 1a and 1b focused on judgment under ignorance or near-ignorance situations, where implementing evaluative approaches is difficult, thereby highlighting reliance on the ignorance prior. In Study 1a, questions presented in the case-prime condition elicited higher ignorance priors compared to the class-prime condition. To address an alternative interpretation suggesting that class formulations might depress probability judgments relative to case formulations, Study 1b involved higher class-primed ignorance priors than case-primed ignorance priors.

Studies 2 and 3 investigated partition dependence in judgment under uncertainty, where individuals can draw on relevant knowledge and may rely on both evaluative assessment and the ignorance prior. Study 2 tasked participants with estimating their score range within a group, leveraging their knowledge of other group members. In Study 3, participants estimated the probability of a candidate’s success in a job application after reading a recommendation letter excerpt of that candidate.

While Studies 1a, 1b, 2, and 3 centered on judgments, Study 4 explored partition dependence in decision-making. Participants were presented with two options and were informed that some would receive a real reward in accordance with their choices. The first option did not involve probability estimation and remained constant across the two conditions. The second option was riskier than the first option. It was formulated with either a case-primed or a class-primed ignorance prior in either the case or the class condition (see Table 7). Participants’ tendency to choose the second option implicitly indicated their probability estimation of the event described by the second option.

We focused our investigation on the studies examining how partition dependence affects judgements and decision making under ignorance and uncertainty: Studies 1a, 1b, 3, and 4 (Table 1). We chose not to replicate Study 2 given that it was conducted in an in-class setting with almost no information provided about both methods, procedure, and results.

## Fox and Rottenstreich (2003): Hypotheses and findings

We did our best to analyze the brief details provided by Fox and Rottenstreich (2003) and summarized our interpretation of the different designs in Table 1, their hypotheses and findings in Table 2, and their descriptives and statistical results in Table 3.

###### Table 1 *Studies 1a, 1b, 3 and 4: Comparison of design*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | 1a | 1b | 3 | 4 |
| Situation | Ignorance | Ignorance | Uncertainty | Ignorance |
| Task | Probability judgment | Probability judgment | Probability judgment | Decision  making |
| Dependent variable | Numerical probability | Numerical probability | Numerical probability | Categorical decision outcome |

*Note*. In situations of ignorance, participants have no or little relevant knowledge. In situations of uncertainty, participants have the opportunity to apply relevant knowledge to some extent.

## 

###### Table 2 *Studies 1a, 1b, 3, and 4: Summary of hypotheses and findings*

|  |  |  |  |
| --- | --- | --- | --- |
| **S** | **Scenario** | **Hypothesis** | **Findings of the target article** |
| 1a | Judgments under ignorance | Individuals utilize partition-dependent probabilities when forming judgments under ignorance when the case prime facilitates a greater ignorance prior of an event’s occurrence than the class prime (1/2 > 1/n). | Participants showed a bias towards partition-dependent ignorance priors. Specifically, participants' judgments differed between the case-prime and the class-prime condition:  (1) Responses were higher in the case-prime condition with an ignorance prior of 1/2 than in the class-prime condition with an ignorance prior of 1/n, indicated by medians of judged probabilities.  (2) 1/2 responses were more common under the case prime than under the class prime. 1/n responses were more common under the class prime than under the case prime. |
| 1b | Judgments under ignorance | Individuals utilize partition-dependent probabilities when forming judgments under ignorance when the class prime facilitates a higher default probability than the case prime (1/2 < x/n). | Participants showed a bias towards partition-dependent ignorance priors. Specifically, participants' judgments differed between the case-prime and the class-prime condition:  (1) Responses were lower in the case-prime condition with an ignorance prior of 1/2 than in the class-prime condition with an ignorance prior of x/n, indicating by both means and medians of judged probabilities.  (2) x/n responses were more common under the class prime than under the case prime. |
| 3 | Judgements under uncertainty | Individuals utilize partition-dependent probabilities when forming judgments under uncertainty (1/2 > 1/n). | Participants showed a bias towards partition-dependent ignorance priors. Specifically, participants' judgments differed between the case-prime and the class-prime condition:  (1) Responses were higher in the case-prime condition with an ignorance prior of 1/2 than in the class-prime condition with an ignorance prior of 1/n, indicated by comparison of medians of judged probabilities.  (2)1/n responses were more common under the class prime than under the case prime. |
| 4 | Decisions under ignorance | Individuals utilize partition-dependent probabilities when making decisions under ignorance (1/2 > 1/n). | Participants showed a bias towards partition-dependent ignorance priors. Specifically, under the case prime, the risker choice had an ignorance prior of 1/2, while under the class prime, the risker choice had an ignorance prior of 1/n. A larger proportion of participants in the case-prime condition made decisions favoring the risker choice when compared to those in the class-prime condition, 23% vs. 11%, *z* =1.95, *p* < .05, one-tailed. |

*Note*. Under case prime, the ignorance prior of the target event’s occurrence is 1/2. In Studies 1a, 3, and 4, under class prime, the ignorance prior of the target event’s occurrence is 1/n where n stands for the number of all possible events in the question. In Study 1b, participants were tasked with estimating the probability of a collection of single events, with x denoting the number of events in that collection. In Study 1a, the comparison between the proportions of 1/2 and 1/n responses relied solely on proportion figures. No statistical test results were provided for this comparison. In Studies 1b, 3, and 4, the χ2 test was used to compare the proportions of 1/n or x/n judgments between the class-prime and case-prime conditions.

###### Table 3 *Studies 1a, 1b, 3 and 4: Summary of reported results*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | Study 1a | Study 1b | Study 3 | Study 4 |
| Dependent variable | | Probability estimation | Probability estimation | Probability estimation | Choice |
| *n* | Case | 41 | 22 | 32 | 74 |
| Class | 53 | 20 | 41 | 70 |
| *Med* | Case | - | 0.69 | 0.40 |  |
| Class | - | 0.71 | 0.23 |  |
| *M* | Case | - | 0.61 | - |  |
| Class | - | 0.72 | - |  |
| Proportion of case-primed 1/2 responses | Case | 19% | - | - |  |
| Class | 8% | - | - | Taking the bet: |
| Proportion of class-primed 1/n (x/n) responses | Case | 19% | 23% | 6% | 23% |
| Class | 49% | 55% | 28% | 11% |
| Mann-Whitney *U* | *U*-statistic | - | - | - |  |
| *p* | < .01 | < .05 | < .05 |  |
| Welch’s *t* | *t* | - | 2.75 | - |  |
| *df* | - | 22 | - |  |
| *p* | - | = .01 | - |  |
|  | *d* | 0.55 [0.05, 0.87] | 0.87 [0.22, 1.51] | 0.47 [0.00, 0.93] |  |
|  | *N* required | 174 | 71 | 238 |  |
| χ2 | χ2 | - | 4.63 | 5.67 |  |
| *df* | - | 1 | 1 |  |
| *p* | - | < .05 | < .05 |  |
| *z* |  |  |  | 1.95 |
|  | *p* |  |  |  | .05 |
|  | *h* | 0.65 [0.13, 0.96] | 0.67 [0.39, 0.95] | 0.62 [0.34, 0.90] | -0.32 [-0.60, -0.05] |
|  | *N* required | 136 | 121 | 164 | 508 |

*Note.* *Med* indicates median. *M* indicates mean. "-" denotes that the results were not reported in the target article. The range of probability estimation spans from 0 (0%) to 1 (100%). For Studies 1a and 3, Cohen’s *d* was estimated from the p values of the Mann-Whitney *U* test using the esc package (Lüdecke, 2019), despite these analyses meant for *t*-test *p*-values, as we assumed that .01 is a good enough proxy for a signal regardless of analysis. For Study 1b, Cohen’s *d* was calculated from Welch’s *t* and using the effectsize package (Ben-Shachar et al., 2020). See accompanying Rmarkdown for effect size calculations.

In Fox and Rottenstreich (2003)’s studies, partition priming was manipulated by the linguistic reformulation of a probability query. In what they referred to as a “case partition”, the target event either will or will not occur. For example, the question “What is the probability that Sunday will be hotter than every other day next week?” would facilitate a binary case partition: Sunday either will or will not be hotter than every other day. In what they referred to as a “class partition”, the probability of a target event is compared against an entire class of possible events. For example, the question “What is the probability that the hottest day of the week will be Sunday?” would facilitate a seven-fold class partition: {Sunday hottest, Monday hottest, . . ., Saturday hottest}. Participants were randomly assigned to make probability estimations in either the case-prime condition, where all the probability questions featured case partition, or the case class-prime condition, where all the probability questions featured class partition. The only distinction between the two experimental conditions was in its linguistic formulation, the target events remained consistent in nature across both conditions.

Given the reporting standards at the time, the reporting of both methodology and the findings were brief with many missing details. In Study 1a, three pairs of questions were presented in each condition. They reported a Mann-Whitney *U* test for all three items and a result of *p* < .01. In Studies 2b and 3, they compared the class-prime and case-prime conditions proportions of judgments that were precisely in accordance with the class-prime ignorance prior (1/n; χ2 test).

## Extensions: Complementary hypotheses

We aimed to extend the replication study by incorporating pairs of complementary hypotheses into Studies 1b and 3. Complementary hypotheses predict mutually exclusive outcomes for the same event. For instance, "H1: Sunday will be the hottest next week" and "H2: Sunday will not be the hottest next week" constitute a pair of complementary hypotheses. The target article only included two complementary hypotheses for each item in Study 1a, but not in Studies 1b and 3. Fox and Rottenstreich (2003) did not elucidate the rationale behind adding the alternative hypothesis. Furthermore, it was unclear whether each participant encountered both mutually exclusive hypotheses or if half of the participants in each condition estimated the probability of a target event occurring (H1), while the other half estimated the probability of the positive target not occurring (H2). Our best guess based on the description “assigned probabilities to three pairs of complementary hypotheses:” is that participants rated both hypotheses.

We therefore decided to expand the original design by incorporating pairs of complementary hypotheses also into Studies 1b and 3. This expansion offers several important advantages. First, it allows for a more comprehensive comparative analysis of the partition priming effects of H1 and H2 responses across the three estimation studies - Studies 1a, 1b, and 3. Through this approach, we can explore whether participants demonstrate the same pattern of reliance on ignorance priors and partition dependence when estimating negative complementary events as their mutually exclusive positive events, and for several scenarios. Second, seeing both hypotheses helps increase the likelihood that participants are processing a fuller range of events, and allows us to indirectly test for both attentiveness and understanding, given that the two probabilities should sum to 100 or close (if there are rounding issues). We contemplated whether to force the two estimations to have to add up to 100, yet decided to instead allow participants to enter whatever approximations to allow for the possibility that participants are systematically deviating from our expected use of the complementary hypotheses, which we could then analyze using exploratory analyses.

## Pre-registration and open-science

We provided all materials, data, and code on: <https://osf.io/g9czs/>. This project received Peer Community in Registered Report Stage 1 in-principle acceptance ((Enter link after IPA); (Enter link after IPA)) after which we created a frozen pre-registration version of the entire Stage 1 packet (Enter link after IPA) and proceeded to data collection. All measures, manipulations, exclusions conducted for this investigation are reported, and data collection was completed before analyses. This Registered Report was written based on the Registered Report template by Feldman (2023).

# Method

[IMPORTANT: Method and results were written using a randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection. For the purpose of the simulation, we wrote things in past tense, but no pre-registration or data collection took place yet.]

## Power and sensitivity analyses

We first calculated effect sizes and conducted a power analysis based on the effects reported in the target article. We used information from the target article (summarized in Table 3) to calculate effect sizes and confidence intervals using R 4.3.3 (R Core Team, 2024) with the help of a guide by Jané et al. (2024). We used the effectsize package (Ben-Shachar et al., 2020) the esc package (Lüdecke, 2019), and pwrss (Bulus, 2023) to compute Cohen’s *d* and Cohen’s *h*.

We concluded that the minimum required sample size was 508 participants in total.‎ We provided more information regarding these calculations in the “Power analysis of the original study effect to assess the required sample for replication” subsection of the supplementary materials.

Moreover, given the likelihood that the target article’s effects are overestimated and the difficulty in computing the effect sizes of Studies 1a, 1b, and 3, we used the “small-telescope” approach (Simonsohn, 2015) aiming for enough power to detect effects much weaker than those reported by the original study (*d*33%) with the general rule of thumb to multiply the largest sample in the target by 2.5. This resulted in a sample of 360 (2.5 times the sample size of Study 4, 144). Given that 360 is less than 508, we maintained the sample size at 508. Accounting for our integrated design, and allowing for the potential of additional analyses, we aimed for a larger total sample of 600 participants, more than 4 times larger than any of the samples in the target article.

We conducted a sensitivity analysis using Gpower (Faul et al., 2007) indicated that a sample of 600 would allow the detection of effect sizes of Cohen’s *d =* 0.27 for independent *t*-tests in Studies 1a, 1b, and 3, and Cohen’s *w* = 0.14 for chisquare to compare proportions reported in Studies 1a, 1b, 3, and 4 (all alpha = 5%, power = 95%, one-tail). These results correspond to weak to medium effects in social psychology Jané et al. (2024). We also note that we completed a similar well-powered Replication Registered Report project conducted with Peer Community in Registered Reports (PCIRR) on diversification bias and partition dependence (Li & Feldman, 2021), with very large effects overall but with one small effect of *d* = 0.27, which we are powered to detect with our target sample.

## Participants

[To demonstrate what the results would look like after data collection we simulated a dataset of 600 participants using Qualtrics and reported our analyses below based on that dataset. Results will later be updated in full to a sample of 600 and the real data.]

We recruited a total of 600 participants from Prolific (*Mage* = XX.XX, *SD* = XX.XX; XXX females, XXX males, XXX other or did not disclose). We note that [XXX] subjects began the survey but [XXX] did not proceed beyond the consent and verifications. We summarized a comparison of the target article sample and the replication samples in Table 4.

[Stage 1 note: We will first pretest the survey duration and technical feedback with 30 participants to make sure our time run estimate was accurate and adjusted pay as needed, the data of the 30 participants will not be analyzed other than to assess survey completion duration, feedback regarding possible technical issues and payment, and needed pay adjustments. Unless in the case of serious technical issues that affect data quality and require survey modification, these participants will be included in the overall analyses.]

[The assignment pay is based on the federal wage of 7.25USD/hour, per minute, so for example 5-8 minutes survey would be paid 1 USD per participant. We first pretested survey duration with 30 participants to make sure our time run estimate was accurate and adjusted pay as needed, the data of the 30 participants was not analyzed other than to assess survey completion duration and needed pay adjustments. For those pretest participants, if survey duration was longer than expected, they were paid a bonus as pay adjustment. The pretest participants' responses were included in the final analysis.]

###### 

###### Table 4 *Difference and similarities between original study and replication*

|  |  |  |
| --- | --- | --- |
|  | Fox and Rottenstreich (2003) | US Americans on Prolific |
| Sample size | Study 1a: 94  Study 1b: 42  Study 3: 73  Study 4: 144  Overall: 353 | 566 | |
| Geographic origin | US American | US American | |
| Gender | not reported | [XXX] males, [XXX] females, [XXX] other/did not disclose | |
| Median age (years) | not reported | [XX] | |
| Average age (years) | not reported | [XX.X] | |
| Standard deviation age (years) | not reported | [XX.XX] | |
| Age range (years) | not reported | [XX-XX] | |
| Medium (location) | not reported (potentially in person on a US university campus) | Computer (online) | |
| Compensation | Study 1a and 3: 1 dollar reward  Study 1b: a donation to charity (the value was not mentioned)  Study 4: Random chance of receiving rewards in line with the participant’s choice | Nominal payment | |
| Year | 2003 or earlier | 2024 | |

## 

## 

### Design: Replication and Extension

In the target article, Studies 1a, 1b, 3, and 4 were conducted separately with independent samples. We ran the four studies together in a single unified data collection, with all scenarios from the four studies presented in random order. Participants were first randomly allocated to either the case-prime or the class-prime condition. This unified design combining replications of several studies into a singular data collection was previously tested successfully in many of the replications and extensions conducted by our team (e.g., Petrov et al., 2023; Vonasch et al., 2023; Yeung & Feldman, 2022; Zhu & Feldman, 2023), also with one successful replication Registered Report on partition dependence with no impact of order effects (Li & Feldman, 2024). We believe that this design is especially powerful in addressing concerns about the target sample (e.g., naivety and attentiveness) when some studies replicate successfully whereas others do not, as well as in allowing for drawing inferences about links between the different studies and consistency in participants’ responding to similar decision-making paradigms.

[Note: In case we fail to find support for the target article’s hypotheses, we will test for order effects (order as a moderator) and for effects for each study when it is displayed first. See “data analysis strategy” section.]

We summarized the experimental design in Table 5 and Table 6, and our adjustments to the target article in Table 7. The baseline main effect design mirroring the target article was a simple between-subjects two partition contrast of case prime versus class prime. The fuller design including the within-subject extensions for Studies 1a/b and 3 were a 2 (case prime versus class prime, between) by 2 (H1 versus H2; within) by 5 (all scenarios; within) mixed design. Additionally, the presentation of task items followed randomized orders.

###### Table 5 *Studies 1a, 1b and 3: Replication and extension experimental design [Between]*

|  |  |
| --- | --- |
| **Case prime condition**  Binary case partition: {a target event will happen vs.  a target event will not happen}  Ignorance prior: 1/2 | **Class prime condition**  N-fold class partition {Event 1 will happen,  Event 2 will happen,...  Event n will happen}  n = number of comparable events in a query.  Ignorance prior: 1/n |
| **DV**: **Probability judgements** [replication]  Participants make probability judgments of a target event which includes a pair of complementary hypotheses (H1 and H2)  There are three items in Study 1a:  Item 1 target event: The noontime temperature at O’Hare Airport on Sunday will be higher than other days next week.  Item 2 target event: the University of Michigan will win the scoring title of the Big Ten Conference in women’s lacrosse for the upcoming season.  Item 3 target event: International Business Machines Corporation (IBM)’s stock price will rise by more than any other stock on the Dow Jones Industrial Average (DJIA) tomorrow.  There is one item in Study 1b and 3 respectively.  Study 1b target event: The warmest day of the week next week (afternoon high temperature) will fall on a weekday (Mon-Fri) rather than the weekend (Sat-Sun).  Study 3 target event: K.T. will be offered a job this year by ACME.  H1: Estimation of the probability that a target event will happen:  Scale: 0 = *Target event will not happen*; 1 = *Target event will happen*  H2: Estimation of the probability that a target event will not happen:  Scale: 0 = *Target event will happen*; 1 = *Target event will not happen* | |
|

*Note*. An ignorance prior is a default judgment which assigns equal likelihood to each comparable event in a query when the judge has little or no relevant knowledge of the events. Study 1a explores judgments under conditions of ignorance, where participants have limited or no access to relevant knowledge.

Study 1b investigated judgments under conditions of ignorance, whereas Study 3 explored judgments under uncertainty. In situations of uncertainty, participants have the opportunity to apply relevant knowledge and may employ a combination of evaluative strategies and ignorance priors. In Study 3, participants are informed that ACME Corporation plans to extend job offers to 10 out of 100 applicants. Subsequently, they read excerpts from a recommendation letter portraying applicant K.T. as cheerful, bright, and hardworking, but somewhat set in her ways.

###### Table 6 *Study 4: Replication and extension experimental design [Between]*

|  |  |
| --- | --- |
| **Case prime condition**  Option 1: Receive $10 for sure.  Option 2: Receive $50 if IBM’s price per share rises by a greater percentage today than any other stock on the DJIA.  Ignorance prior of option 2: 1/2 | **Class prime condition**  Option 1: Receive $10 for sure  Option 2: Receive $50 if the stock whose price per share rises by the greatest percentage on the DJIA today is IBM.  Ignorance prior of option 2: 1/30 |
| **DV**: **Decision** [replication]  Participants make decisions between two options.  Scale (categorical): option 1; option 2 | |
|

*Note.* Study 4 examined decisions under ignorance. All participants were reminded that “the Dow-Jones Industrial Average (DJIA) consists of 30 large industrial stocks, including International Business Machines Corporation (IBM)”.

We had to make several adjustments to the target article’s stimuli. In Study 1a, the original target event of Item 2 identified by Fox and Rottenstreich (2003) was the University of Illinois winning the scoring title. In 2003, the Big Ten Conference in women’s lacrosse consisted of 11 teams. However, in recent years, the league has been reduced to seven teams, with the University of Illinois no longer participating. Consequently, we adjusted this item to estimate the University of Michigan's winning rate. Fox and Rottenstreich (2003) did not specify the Big Ten Conference having 11 teams. They assumed their participants would know that information, and that the ignorance prior for most participants would be 1/10. In contrast, given our broader sample, we needed to explicitly inform our participants that there are seven teams in the league so that they can make an informed decision. Therefore, our ignorance prior for this item was set at 1/7. This is a deviation from the target, but we felt a needed and crucial one because without this information and/or asking participants for their knowledge about that league, we are assuming too much and many things can go wrong.

[Note to reviewers: This is one of several crucial adjustments we had to make to address what we felt were important weaknesses in that target design. However, we can also see challenges in these adjustments, as - for example - incorporating a statement clarifying that there are seven teams competing in the league before presenting the question may bias participants in some way. However, given that the information is presented equally to both conditions so would impact both conditions in a similar way and would hopefully not impact the manipulation, and so we believe the benefits far outweigh the weaknesses. We welcome your opinion and suggestions if you can think of a better way of addressing these or improving our adjustments. ]

Likewise, in both item 3 of Study 1a and Study 4, Fox and Rottenstreich (2003) originally targeted General Motors’ (GM’s) stock price rising more than any other stock on the Dow Jones Industrial Average (DJIA). However, the components of the DJIA have undergone changes over time and GM is no longer a constituent of the DJIA. Consequently, we have replaced GM with International Business Machines Corporation (IBM), which currently stands as a component of the DJIA (2024).

[Note to reviewers: This is one of several crucial adjustments we had to make to account for changes in time and outdated stimuli. We welcome your opinion and suggestions if you can think of improving our adjustments. ]

## Procedure

[*For review: The Qualtrics survey .QSF file and an exported DOCX file are provided on the OSF folder. A preview link of the Qualtrics survey is provided on:*<https://hku.au1.qualtrics.com/jfe/preview/previewId/30327764-186d-47f0-b59e-b24a9df94f42/SV_eLNVJDaGPAyx9MW?Q_CHL=preview&Q_SurveyVersionID=current>

We reconstructed the target’s survey items of Study 1a, 1b, 3 and 4 and adjusted it to an online Qualtrics survey based on the information provided in the article.

Participants were initially presented with the informed consent form and a detailed outline of the study's requirements and procedures. To proceed, they were asked to indicate their consent with four questions confirming their eligibility, understanding, and agreement with study terms, which they must answer with a “yes” and required responses in order to proceed to the study. Three of the four questions also served as attention checks, with the options order being rotated (yes, no, not sure). Failing those attention checks would mean that the participants did not indicate consent and therefore could not embark on the study.

Upon confirming their consent and demonstrating understanding of the study instructions, participants were randomly assigned participants to either the case-prime condition or the class-prime condition (“evenly presented” randomizer in Qualtrics). Within each condition, participants answered all scenarios from the four studies of the target article in a randomized order.

We implemented two questions of competing hypotheses for all scenarios in Studies 1a, 1b, and 3 (see clarification above). We note that in Study 1a of the target article, it was reported that participants encountered three pairs of complementary hypotheses, yet Fox and Rottenstreich (2003) did not specify whether each pair was presented concurrently or if all six questions (3 items × 2 hypotheses) were presented in a counterbalanced manner. Given the repetition of the complementary hypotheses in all five scenarios, we were concerned that randomizing the order would confuse participants, and so presented the pairs in fixed order, where the “will” hypothesis is always presented first, and the “will NOT” hypothesis is always presented second.

In both conditions, participants made a series of probability prediction queries including several numerical probability estimations and a two-option decision-making task. Under the case prime, the queries were designed to facilitate a binary case partition with an ignorance prior of 1/2. Under the class prime, the queries were designed to facilitate a n-fold class partition with an ignorance prior of 1/n. At the end of the experiment, participants answered a number of funneling and demographic questions.

## Manipulations

### Partition priming (between-groups)

Each participant was randomly assigned to complete tasks involving likelihood estimation either in the case-prime or the class-prime condition with either case or class formulations of the same queries.

### Scenario (within-groups)

Each participant completed two types of tasks: numerical probability estimation task and two-option decision making task. For the numerical probability estimation task, participants rated the likelihood of an event’s occurrence in percentages. In the decision-making task, the first option was consistent across both conditions, whereas the second option included an ignorance prior aligned with either case or class formulations of the same event. Participants were directed to select between the two options. Although participants' tendency to select the second option correlated with their probability estimation of the occurrence of the target event associated with the second option, they were not required to provide numerical responses during the decision-making task.

## Measures

### Replication

#### Studies 1a, 1b and 3: Probability estimation

Based on the information presented in the target article, the participants in Fox and Rottenstreich's study (2003) recorded probability estimations in both fractions and decimals (e.g., 1/2 and 0.5). In our study, we maintained the probability scale from 0 to 1 as used in the target article, but we specifically instructed participants to record probabilities as percentages. The percentage scale ranged from 0% to 100%, with 0% indicating that the target event will not happen and 100% indicating that the target event will happen (see Table 5).

#### Study 4: Choice

The decision outcome is either option 1 or option 2 (see Table 6).

## 

###### Table 7 *Replication and extension adjustments to the target article’s methods and design*

| **Study** | **Factor** | **Target article** | **Adjustment in current study** | **Reason for change / Justifications** |
| --- | --- | --- | --- | --- |
| 1a | Item 2 (see Table 6 for details) | The target event was the University of Illinois winning the Big Ten Conference in women’s lacrosse | We changed the target event into the University of Michigan winning the Big Ten Conference in women’s lacrosse | The University of Illinois is not currently competing in the Big Ten Conference in women’s lacrosse. |
| 1a | Item 2 ignorance prior and information presentation | 1/10 | 1/7  Explicitly informing participants regarding the number of teams (7). | Fox and Rottenstreich (2003) did not mention to the participants that the Big Ten Conference comprised 11 teams. Given the relative obscurity of the league and its suggestive name, they assumed that most participants would perceive the league as having 10 teams. |
| 1a | Item 3 (see Table 5 for details) | The target event was GM’s stock price rising by the greatest amount on the DJIA | We changed the target event into IBM’s stock price rising by the greatest amount on the DJIA | GM is not a component of the DJIA currently (2024). |
| 1b, 3, | Extra items | Only Study 1a incorporated pairs of complementary hypotheses of the same event. | We added a complementary hypothesis to each item in Study 1b and 3. | Refer to the “extensions” section for details. |
| 1a, 1b, 3 | Data format: probability | Participants recorded probability estimations in both fractions and decimals | We specifically asked our participants to record probability estimation in percentages, fraction answers were not allowed. We clarified in the instructions that participants could convert a fraction response into a percentage. | Consistency, ensuring same processing mode across participants, allowing for answer validation in Qualtrics, and for more accurate data analysis that does not require conversions. |
| 1a, 1b, 3 | Classification of responses align with ignorance priors | ±1% or exact values of the ignorance prior | ±5% of the ignorance prior | Consistency across Studies 1a, 1b, and 3, |
| 1a | Data analysis | Mann-Whitney *U* test | Mann-Whitney *U* test Welch’s *t* test  Chi squared test (the proportion of case-primed responses and class-primed responses in the two conditions) | Increase transparency and improve on the target’s data analysis. |
| 1b | Data analysis | Mann-Whitney *U* test Welch’s *t* test  Chi squared test (the proportion of class-primed responses in the two conditions) | Mann-Whitney *U* test  Welch’s *t* test  Chi squared test (the proportion of case-primed responses and class-primed responses in the two conditions) | Increase transparency and improve on the target’s data analysis. |
| 3 | Data analysis | Chi squared test (the proportion of class-primed responses in the two conditions) | Mann-Whitney *U* test  Welch’s *t* test  Chi squared test (the proportion of case-primed responses and class-primed responses in the two conditions) | To enhance transparency and refine the data analysis strategy. |
| 1a (item 3) and 4 | Data analysis | There was no analysis for order effect because Studies 1a and 4 were conducted independently with different participants. | We will test the order effect. Specifically, we will examine whether the order of task presentation (Study 1a Item3 and Study 4) influence decision-making outcomes in Study 4. | To explore whether the reliance on ignorance priors is influenced by the use of a numerical response scale or not. |
| 1a, 1b, and 3 | Data analysis | Mann-Whitney *U* test,  Welch’s *t* test, and Chi squared test. | An additional 2 (condition: between-subjects) × 5 (question: within-subjects) × 2 (event: within-subjects) mixed ANOVA. | A mixed ANOVA allows for the examination of interaction effects between condition (between-subjects), and question and event (within-subjects), providing a comprehensive analysis of how these factors influence participants’ responses |
| 1a, 1b, 3, and 4 | Compensation | Study 1a and 3: 1 dollar reward  Study 1b: a donation to charity (the value was not mentioned)  Study 4: Random chance of receiving rewards in line with the participant’s choice | 7.25+USD/hour across the four studies | The original studies exclusively targeted university students, affording Fox and Rottenstreich (2003) greater flexibility in compensation adjustment. Our participant recruitment will be conducted through Prolific, which mandates a minimal hourly reward requirement. Consequently, we have opted not to replicate the initial payments adjusted for inflation. |
| 4 | Study 4 Reward | Participants were informed that randomly selected respondents would receive actual reward based on their choices. | We did not randomly select participants for actual reward in accordance with their choice. | To allow for comparison to the other hypothetical studies. Control costs, and reduce the complexity in experiment implementation. |

## 

## 

## Evaluation criteria for replication findings

We pre-registered our strategy to evaluate our conclusion of whether the target article successfully replicated overall based on the number of studies in which our findings indicated a signal in the same direction as the target article, per the following: successful: three or four out of four studies; a failed replication: no studies; mixed findings: if only one or two studies.

## Replication closeness evaluation

We provided details on the classification of the replications using the criteria by LeBel et al. (2018) criteria in Table 8. Given the adjustments to the stimuli, data analysis, our extensions, and the unified design, we summarized the replication as being between direct and conceptual, a “close to far” replication. Much was based on the target article, but any replication that would aim to repeat the target’s method would need to make adjustments and our extensions and planned analyses overall help strengthen the replication, improve accuracy, and reduce noise.

###### Table 8 *Classification of the replication, based on LeBel et al. (2018)*

|  |  |  |
| --- | --- | --- |
| **Design facet** | **Replication** | **Details of deviation and severity [minor/major]** |
| Effect/hypothesis | Same |  |
| IV construct | Same |  |
| DV construct | Same |  |
| IV operationalization | Similar | We made minor adjustments to the scenarios to update them to current times (see Table 7). |
| DV operationalization | Different | Major: Across all studies we elicited both the affirmative and the negative of the probability (will versus will not occur). |
| IV stimuli | Same |  |
| Procedural details | Different | Major;  Procedure: Fox and Rottenstreich (2003)’ recruited different participants for Studies 1a, 1b, 3, and 4. We combined the four studies into a cohesive single data collection. Each participant completed tasks from all four studies.  Compensation: Fox and Rottenstreich (2003) employed varying compensation schemes across the four studies we chose to replicate, as detailed in Table 4. We have adhered to the standard payment rates on Prolific. |
| Contextual variables | Same |  |
| Population (e.g., age) | Different | Major; Fox and Rottenstreich (2003)’s participants were exclusively undergraduate and M.B.A students of US universities. Our participants were drawn from the general US population. |
| Replication classification | Close to far replication |  |

*Note*. Criteria for evaluation of replications by LeBel et al. (2018). "Similar" category was added to the LeBel et al. (2018) typology to refer to minor deviations or extensions aimed to adjust the study to the target sample that are not expected to have major implications on replication success.

## Data analysis strategy

### Replication: As in the original

To mirror the target article’s analyses, for Studies 1a, 1b, and 3, we first ran the Mann-Whitney *U* test and Welch’s *t*-test to compare the probability judgements between the case-prime condition and the class-prime condition. We then ran the Chi-squared test to compare the prevalence of responses aligning with ignorance priors in each priming condition. Deviating from the target, we will base our assessment of overall support for the hypotheses based on the overall means, rather than the comparison of the arbitrary categorization of match with ignorance priors. We feel that the theory of partition dependence is more aligned with the overall shift in preference rather than whether some participants were closer to ignorance priors, yet we will supplement our reporting with chi-square for the sake of consistency with the target and completeness.

There were three items in Study 1a. We conducted a series of Chi-squared tests to assess the proportions of case-primed and class-primed responses for each item. Furthermore, we compared the overall proportion of case-primed and class-primed responses across all three items. However, we did not compare the overall difference in probability estimation of the three items between the two conditions using Mann-Whitney *U* test and Welch’s *t*-test.

For Study 4, we ran a two proportion *z*-test (/chi square) to compare the difference in option proportions between the two priming conditions.

### Replication: Additional analyses

We aimed to supplement the target article’s analyses by running the Mann-Whitney *U* test, Welch’s *t*-test, and Chi-squared test for each study of Studies 1a, 1b, and 3 in the target article. We conducted the following analyses for all three studies: 1) Mann-Whitney U test and Welch’s *t*-test to compare the probability judgements between the case-prime condition and the class-prime condition, 2) Chi-squared test to determine the prevalence of responses in each priming condition. Specifically, we examined: (a) proportions of judgments aligned with the case-prime ignorance prior (1/2), and (b) proportions of judgments aligned with the class-prime ignorance prior (1/n or x/n) in the two conditions.

Though not well-specified in the target article, we made the decision to classify all responses falling within ±5% of the ignorance prior as influenced by partition priming. For instance, in Study 1b, the ignorance prior of the case condition was 50%, and responses ranging from 45% to 55% were identified as being case-primed. Similarly, with an ignorance prior of 71% for the class condition, responses falling within the range of 66% to 76% were classified as class-primed. No item had overlapping response ranges. For item 3 in Study 1a, responses ranging from 0% to 8% were classified as being class-primed. In the target article, Fox and Rottenstreich (2003) utilized varying definitions of response ranges for case-primed or class-primed conditions across different items. For instance, some items employed a ±1% range (e.g., 0.14 or 0.15 for item 1 of Study 1a), while others were rounded to two decimal places (e.g., 0.03 for item 3 of Study 1a), and certain items utilized exact numbers (e.g., 0.1 for Study 3).

[Note to reviewers: This was ambiguous in the target article, though it seems like they employed different criteria across studies. In addition, small ranges might pose a problem given that people are often bad or lazy to calculate and accurately report operations like 1/7 = 1 divided by 7 = 0.142. We think a range of ±5% across all studies seems more appropriate, even if deviating from the target. We acknowledge that ±5% also feels somewhat arbitrary, but seems like a good compromise between too small and too large. If you have insights or suggestions, we welcome a better approach.]

### Extensions: Competing hypotheses and scenarios

To analyze our extension, we examined whether participants’ responses to negative events still exhibit a systematic bias towards partition-dependent ignorance priors using a 2 x 2 x 5 mixed ANOVA for Studies 1a/b and 3 contrasting the two competing hypotheses with an interaction of the case versus class conditions, for each of the studies. We generally expect no differences, and have no specific predictions as to the results, and therefore the analysis is exploratory. This approach would serve to examine differences across scenarios and in competing hypotheses framing, as a potential examining of the generalizability of the phenomenon.

### Extension: A three-way mixed ANOVA for Studies 1a, 1b, and 3

In our analysis plan for Studies 1a, 1b, and 3, extending upon the previously outlined methods, we will employ a three-way mixed ANOVA to examine participants' probability judgments: 2 (condition; case vs. class; between-subjects) × 5 (study: Study 1a Item 1, Study 1a Item 2, Study 1a Item 3, Study 1b, Study 3; within-subjects) × 2 (evaluation type: happen vs. will not happen; within-subjects).

In addition, we will aim to interpret the results in reference to how they fit our ignorance prior range.

### Extension: Order effect (Study1a Item3 and Study4)

In the original study, the rationale for Study 4 was to investigate whether the reliance on ignorance priors depended on using a numerical response scale. In Item 3 of Study 1, participants assessed the numerical probability of a stock price rising to the highest level in the DJIA. In contrast, Study 4 required participants to evaluate the probability without providing numerical responses. Fox and Rottenstreich (2003) conducted Studies 1 and 4 independently with different participants. As we plan to conduct all four studies consistently with the same participants, we will explore whether the reliance on ignorance priors is influenced by the use of a numerical response scale by examining the order effect between Item 3 of Study 1a and Study 4. Participants who complete Item 3 of Study 1a before Study 4 will already possess an explicit numerical probability for making decisions in Study 4. Conversely, those who complete Study 1a Item 3 after Study 4 will make decisions in Study 4 without an explicit probability judgment.

To explore how different conditions and the order of task presentation influence decision-making under uncertainty, we will perform a binomial logistic regression analysis. The dependent variable is the binary choice between two distinct financial options. Option 1 entailed receiving $10 with certainty, while Option 2 was contingent upon the price of IBM' stock outperforming all other stocks on the DJIA. The independent variables are order and condition. Order is a within-subjects factor with two levels: "Judgement First," where participants completed numerical estimation initially, and "Decision First," where participants made a non-numerical decision initially. Experimental condition is a between-subjects factor (case vs. class)

### Outliers and exclusions

In this study, we would not classify any cases as outliers. We included all the data collected in our analysis for those who successfully completed the entire study.

### Order effects

One deviation from the target article is that all participants completed all scenarios in random order. We considered this to be a stronger design with many advantages, yet one disadvantage is that answers to one scenario may bias participants’ answers to the following scenarios.

We therefore pre-register that if we fail to find support for the core hypotheses of the target article in any of the scenarios that we rerun exploratory analyses for only the failed scenario by focusing on the participants that completed that study first, and examine order as a moderator. To compensate for multiple comparisons and increased likelihood of capitalizing on chance, we will set the alpha for the additional analyses to a stricter .005.

# Results

[IMPORTANT: Method and results were written using a randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection. For the purpose of the simulation, we wrote things in past tense, but no pre-registration or data collection took place yet.]

We summarized the descriptives in Table 9 and statistical tests in Table10 and 11. The following analyses were performed with JAMOVI 2.4.8 (The JAMOVI project, 2023) and R 4.3.3 (R Core Team, 2024).

## Replication: Studies 1a, 1b, and 3

We summarized descriptive statistics of Studies 1a, 1b, and 3 in Table 9. We conducted a series of Mann-Whitney *U* tests and Welch’s *t*-tests to compare probability estimations in Studies 1a, 1b and 3, plotted in Figure 1.

###### Table 9 *Study 1a, b, and 3: Descriptive statistics of probability estimation of H1 (happen) target events*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Case prime (*n* = 300) | | | | Class prime (*n* = 300) | | | |
| S | Item | Ignorance prior | *Med (*%) | *M (*%) | *SD (*%) | Ignorance prior | *Med (*%) | *M* *(*%) | *SD* *(*%) |
| 1a | Weather | 1/2 (50%) | 46.50 | 49.01 | 29.63 | 1/7 (14%) | 54.00 | 52.43 | 27.97 |
| Sports | 1/2 (50%) | 44.50 | 47.68 | 28.81 | 1/7 (14%) | 51.00 | 50.01 | 27.35 |
| Business | 1/2 (50%) | 47.00 | 48.53 | 28.01 | 1/30 (3%) | 48.50 | 49.18 | 29.31 |
| 1b | Weather | 1/2 (50%) | 45.00 | 48.32 | 29.08 | 5/7 (71%) | 47.00 | 47.69 | 28.15 |
| 3 | Offer | 1/2 (50%) | 50.00 | 49.48 | 28.92 | 1/10 (10%) | 50.00 | 49.77 | 31.13 |

*Note*. *Med* indicates median judged probability, *M* indicates mean, *SD* indicates standard deviation, *n* indicates sample size. The units of *Med, M, and SD* are percentages.

###### 

###### Table 10 *Study 1a, b, and 3: results of Mann-Whitney U and Welch’s t test comparing probability estimation (H1)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Study | | 1a | | | 1b | 3 |
|  |  | Item 1 | Item 2 | Item 3 |
| Mann-Whitney *U* | *U* | 41958.50 | 42827.50 | 44397.50 | 44459.50 | 44712.00 |
|  | *p* | .152 | .306 | .777 | .799 | .892 |
|  | Rank biserial r | .07 | .05 | .01 | .01 | .01 |
| Welch’s *t* | *t* | -1.45 | -1.02 | -0.28 | 0.27 | -0.12 |
|  | *df* | 596.03 | 596.39 | 596.77 | 597.36 | 594.79 |
|  | *p* | .147 | .310 | .779 | .788 | .906 |
|  | Cohen’s *d* and 95% CI | -0.12  [-0.28, 0.04] | -0.08  [-0.24, 0.08] | -0.02  [-0.18, 0.14] | 0.02  [-0.14, 0.18] | -0.01  [-0.17, 0.15] |
|  | log*e*​(*BF*01​) | 1.4 | 1.9 | 2.4 | 2.4 | 2.4 |
|  | posterior difference and 95% CI | -3.3  [-7.9, 1.1] | -2.3  [-6.7, 2.1] | -0.63  [-5.2, 4.0] | 0.62  [-3.9, 5.0] | -0.28  [-5.1, 4.5] |

###### 

###### Table 11 *Studies 1a, b, and 3: proportions of response aligned with ignorance priors (H1)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Condition | Study 1a Q1 | Study 1a Q2 | Study 1a Q3 | Study 1b | Study 3 |
| Case-primed response | Case | 10% | 10% | 9% | 7% | 11% |
| Class | 12% | 10% | 12% | 13% | 11% |
| Class-primed response | Case | 13% | 13% | 9% | 11% | 10% |
| Class | 9% | 12% | 9% | 9% | 12% |

*Note.* Responses falling within ±5% of the ignorance prior were considered to be influenced by partition priming (aligned with ignorance priors).

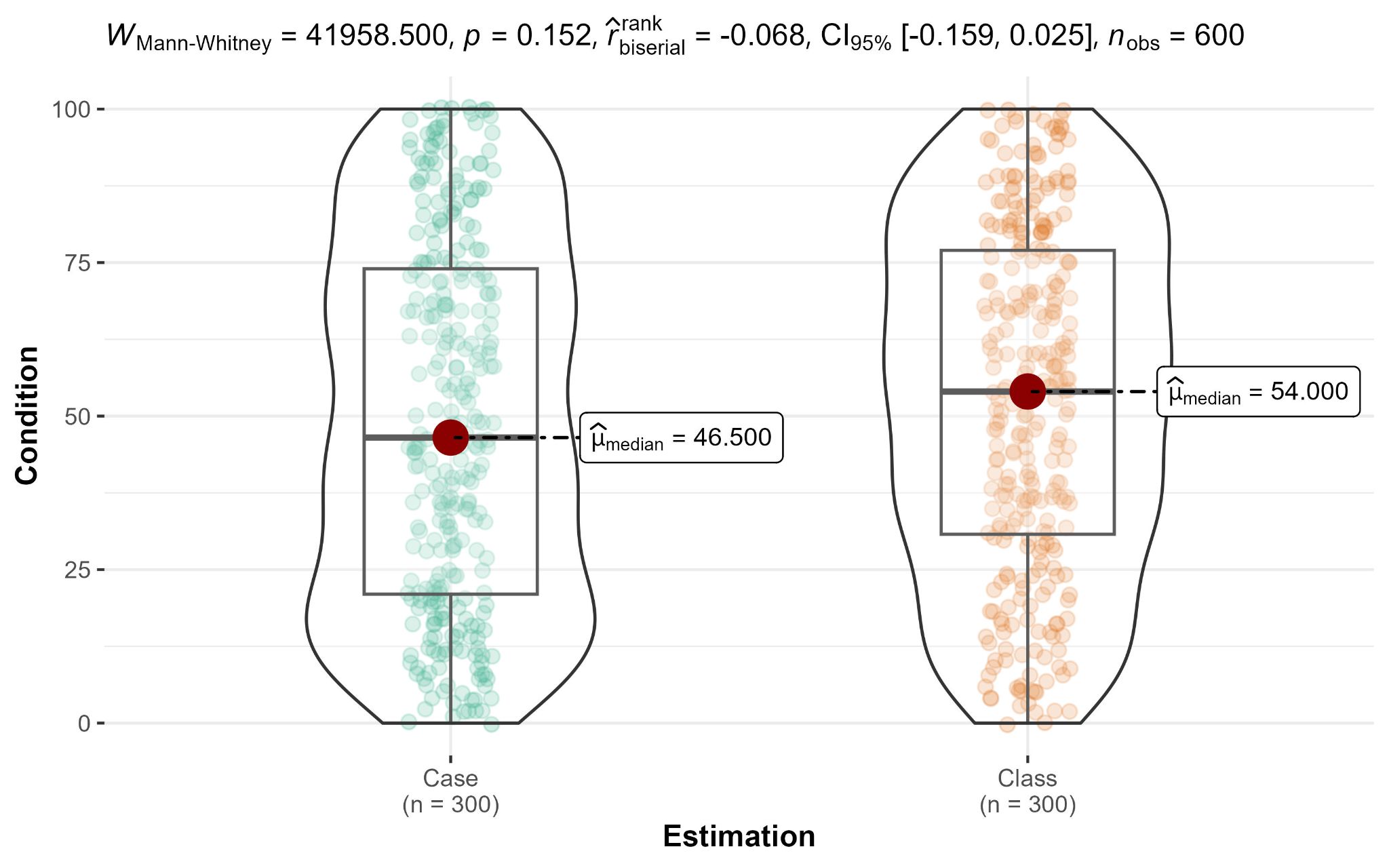
###### Table 12 *Studies 1a, b, and 3:* χ2 *test comparing the proportions of case-primed and class-primed responses in the two experimental conditions (H1)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Study | Case-primed response | Class-primed response |
| χ2 (*df* = 1) | 1a Q1 | 0.27 | 2.39 |
|  | 1a Q2 | 0.02 | 0.14 |
|  | 1a Q3 | 1.80 | 0.00 |
|  | 1b | 5.27 | 0.65 |
|  | 3 | 0.07 | 0.27 |
| *p* | 1a Q1 | .601 | .122 |
|  | 1a Q2 | .893 | .710 |
|  | 1a Q3 | .180 | 1.00 |
|  | 1b | .022 | .421 |
|  | 3 | .794 | .602 |
| Cohen’s *h* and 95%CI | 1a Q1 | -0.07 [-0.32, 0.19] | 0.20 [-0.05, 0.45] |
| 1a Q2 | -0.02 [-0.28, 0.25] | 0.05 [-0.20, 0.29] |
| 1a Q3 | -0.18 [-0.44, 0.08] | 0.00 [-0.28, 0.28] |
| 1b | -0.31 [-0.58, -0.05] | 0.11 [-0.15, 0.37] |
| 3 | 0.03 [-0.22, 0.29] | -0.07 [-0.32, 0.19] |

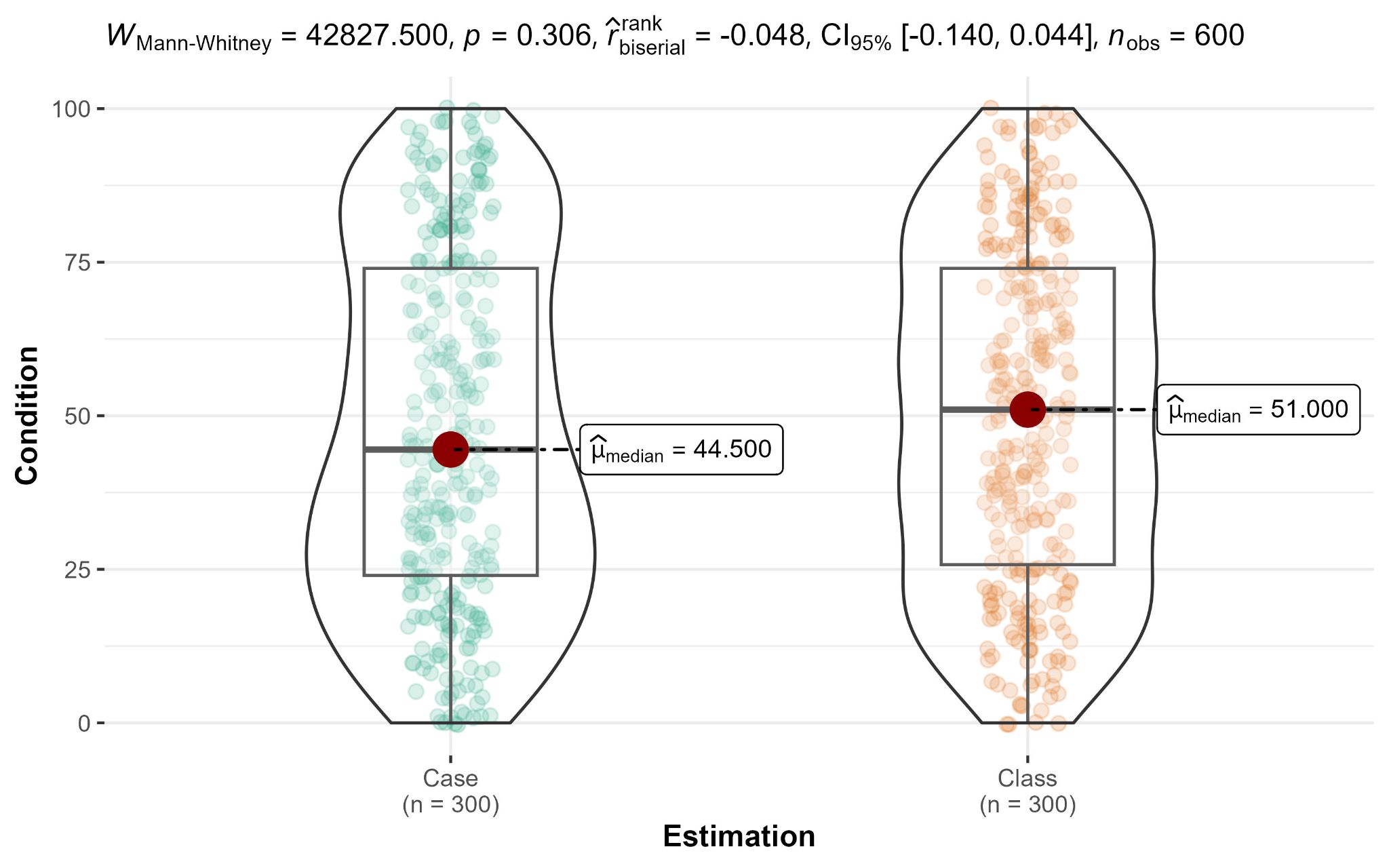
*Note*. *N* = 600

###### Figure 1 *Studies 1a, 1b, and 3: Mann-Whitney tests for estimations*

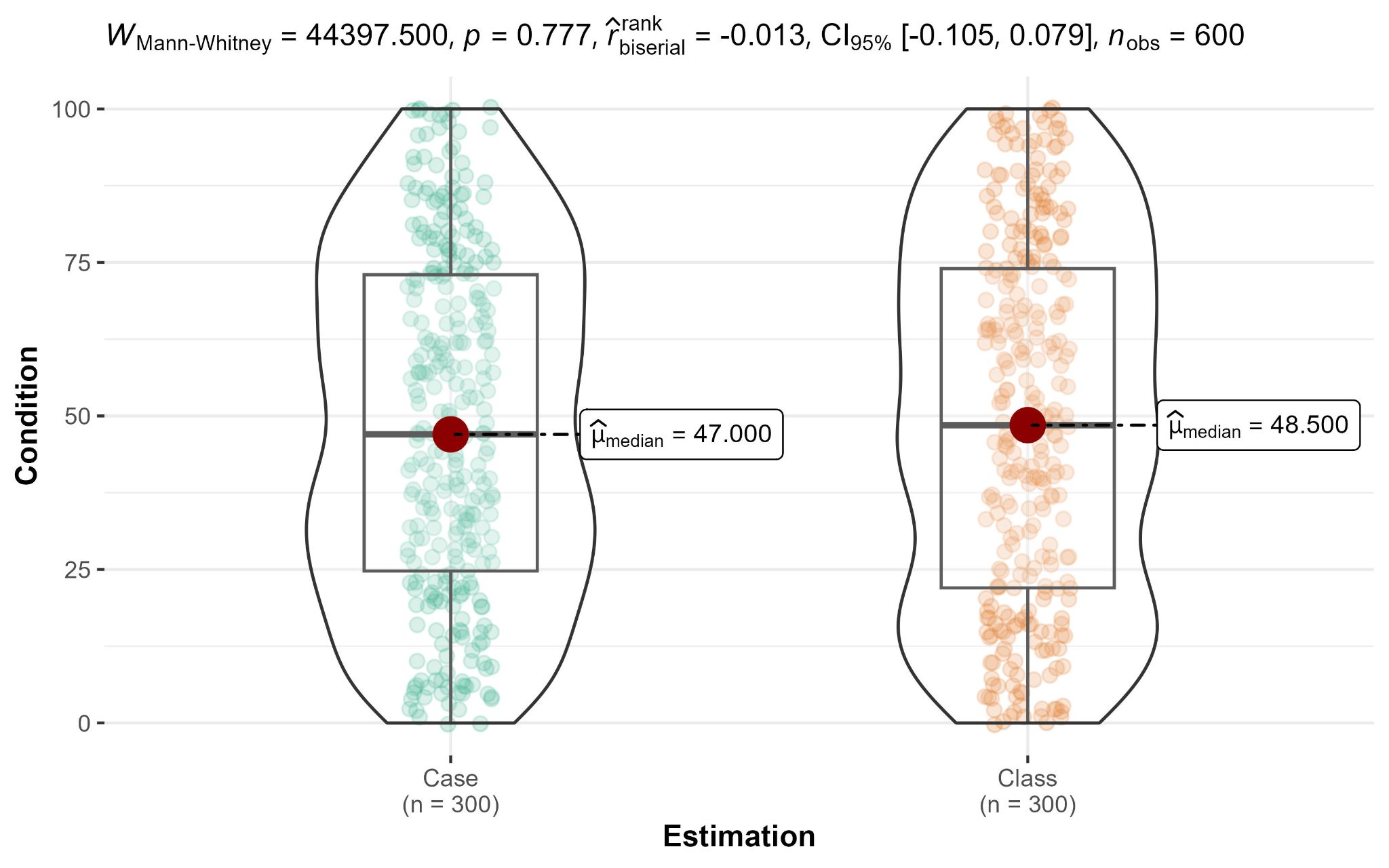
Study 1a Q1



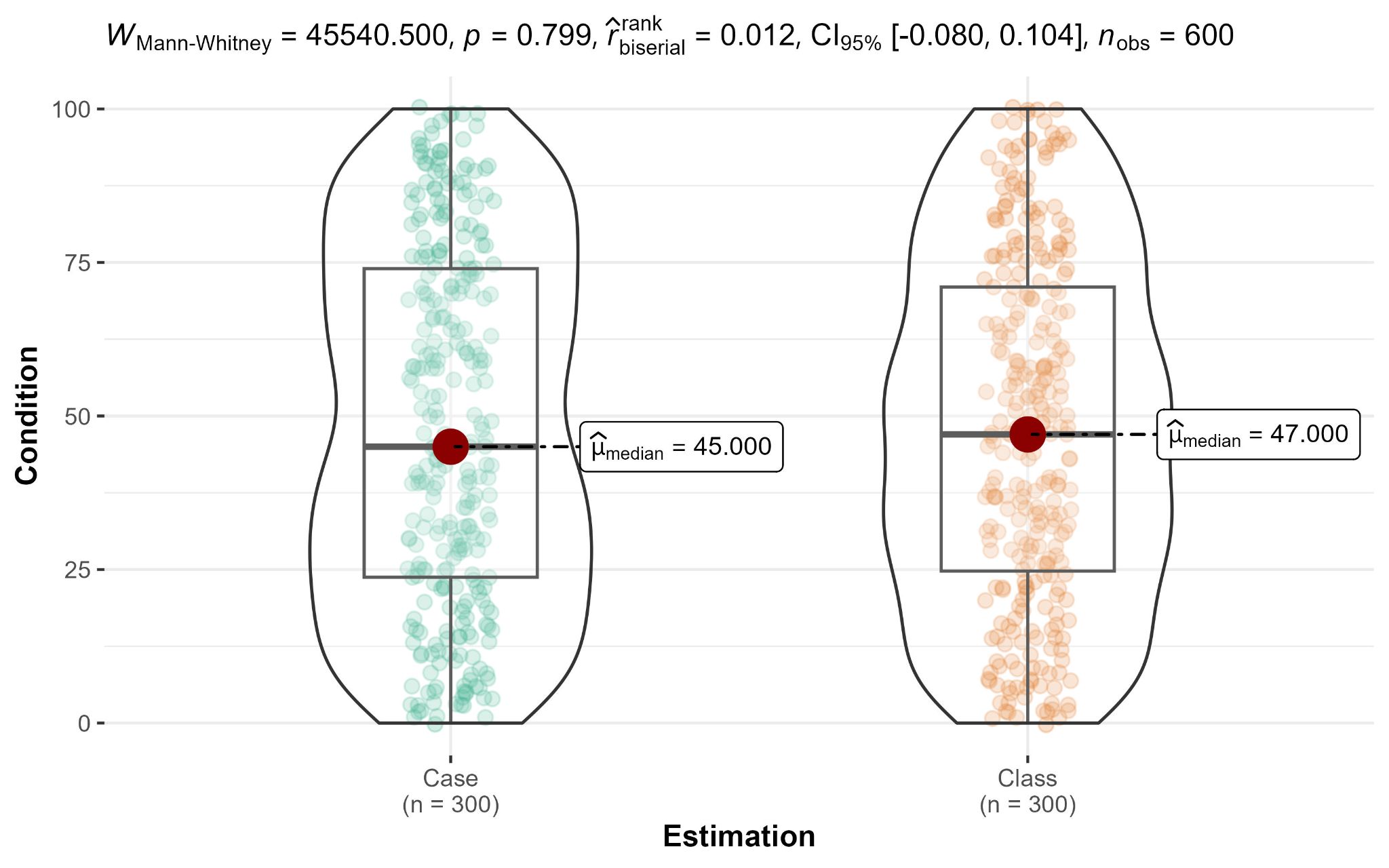
Study 1a Q2



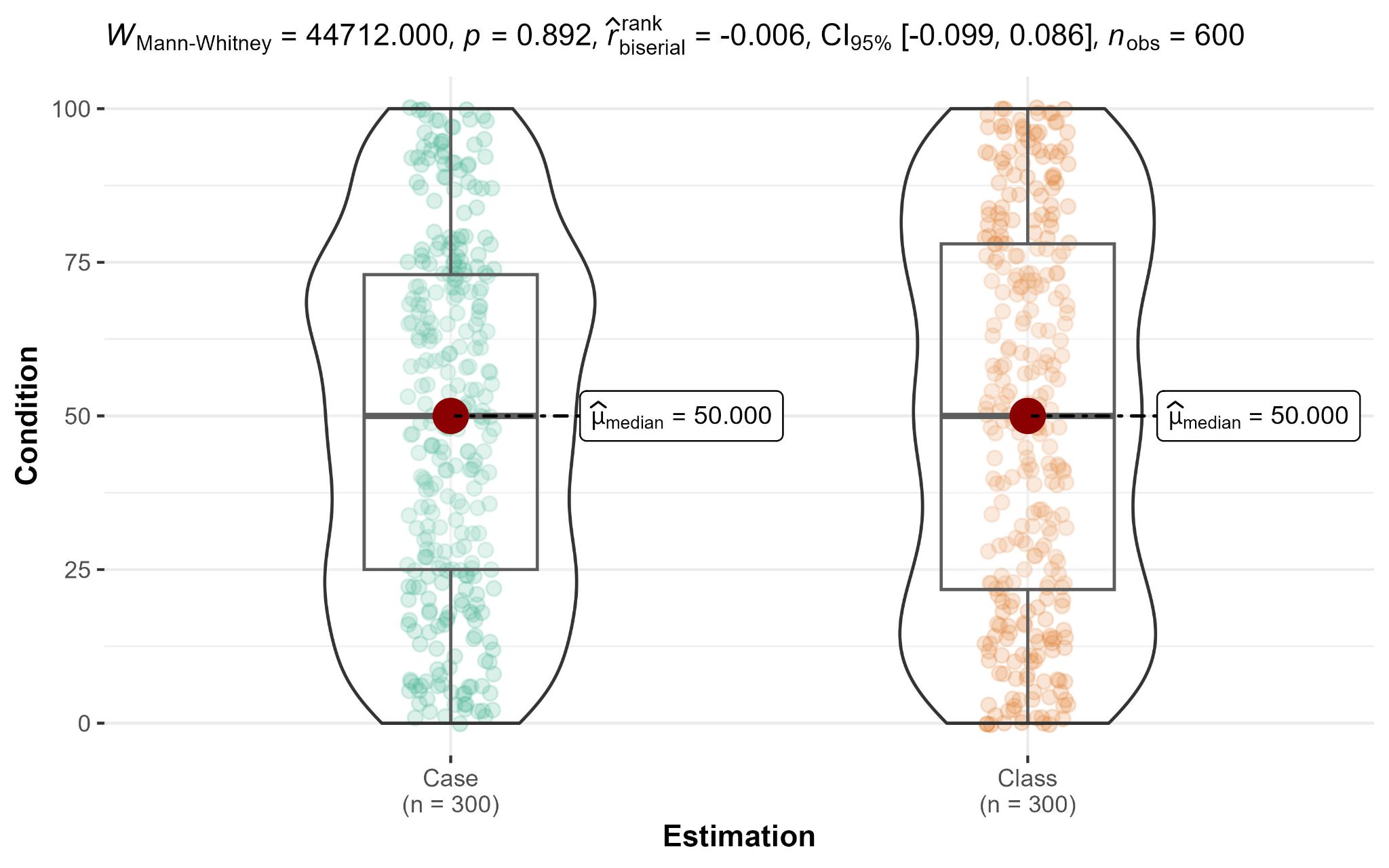
Study 1a Q3



Study 1b

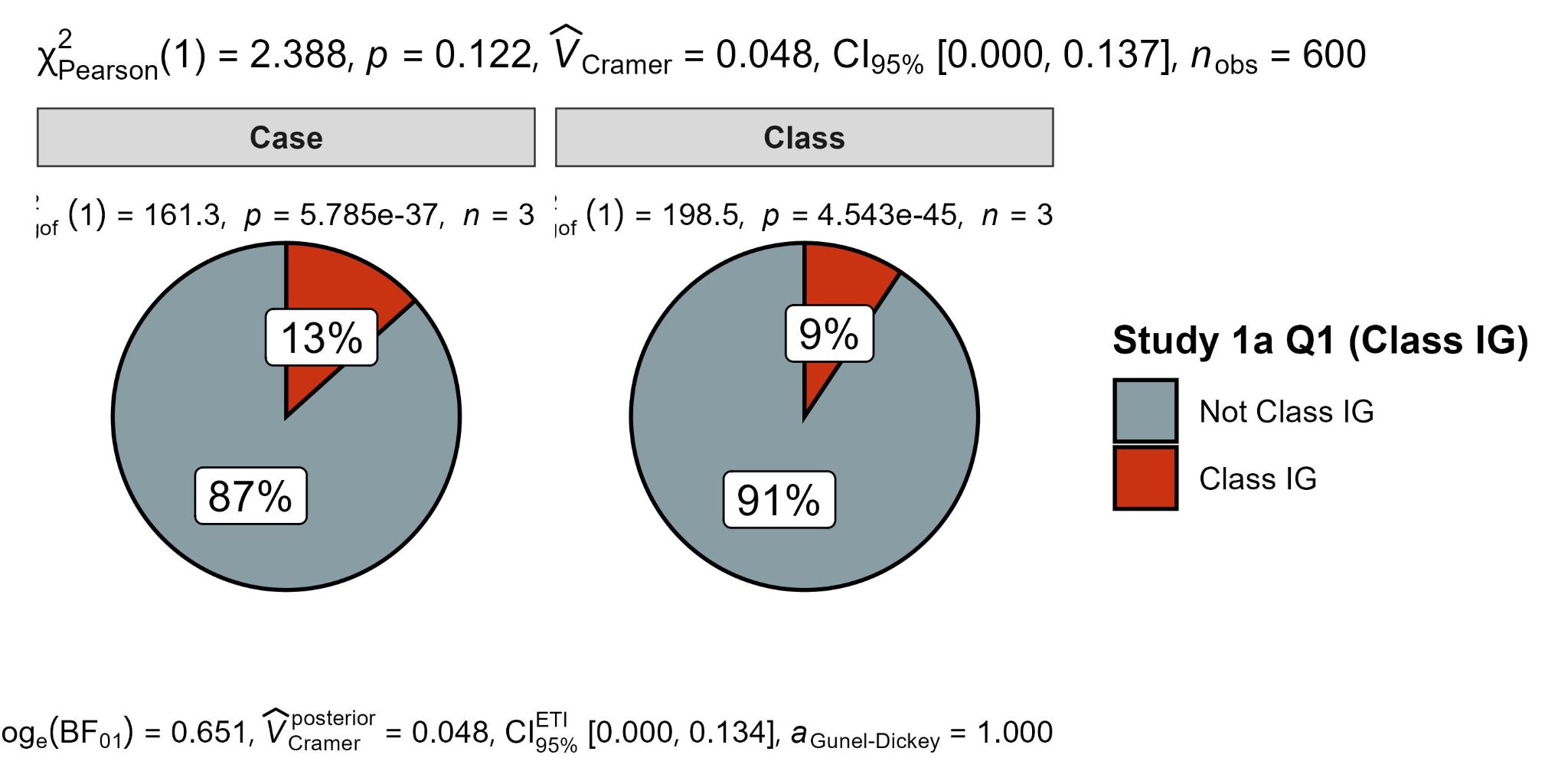
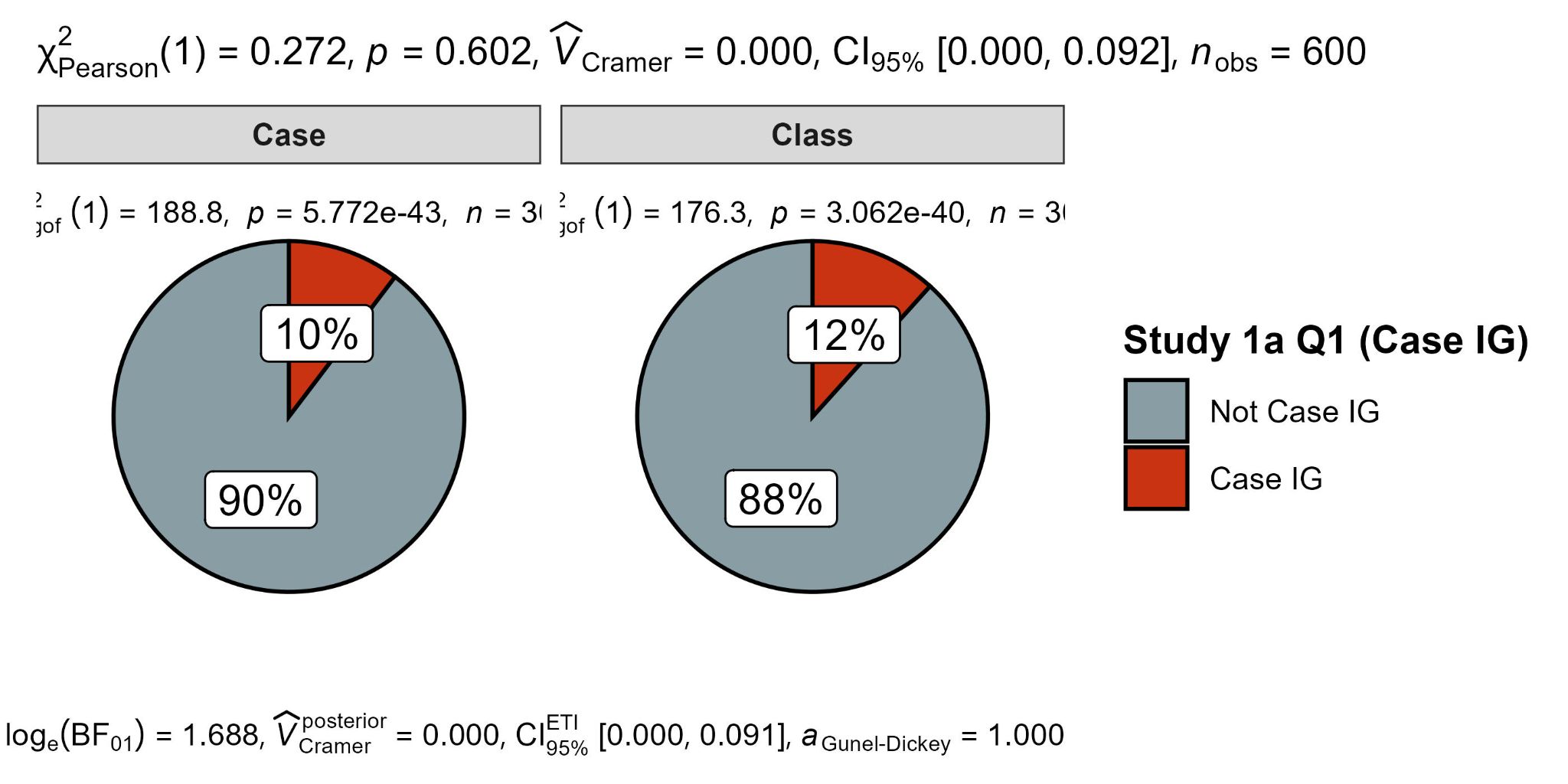


Study 3

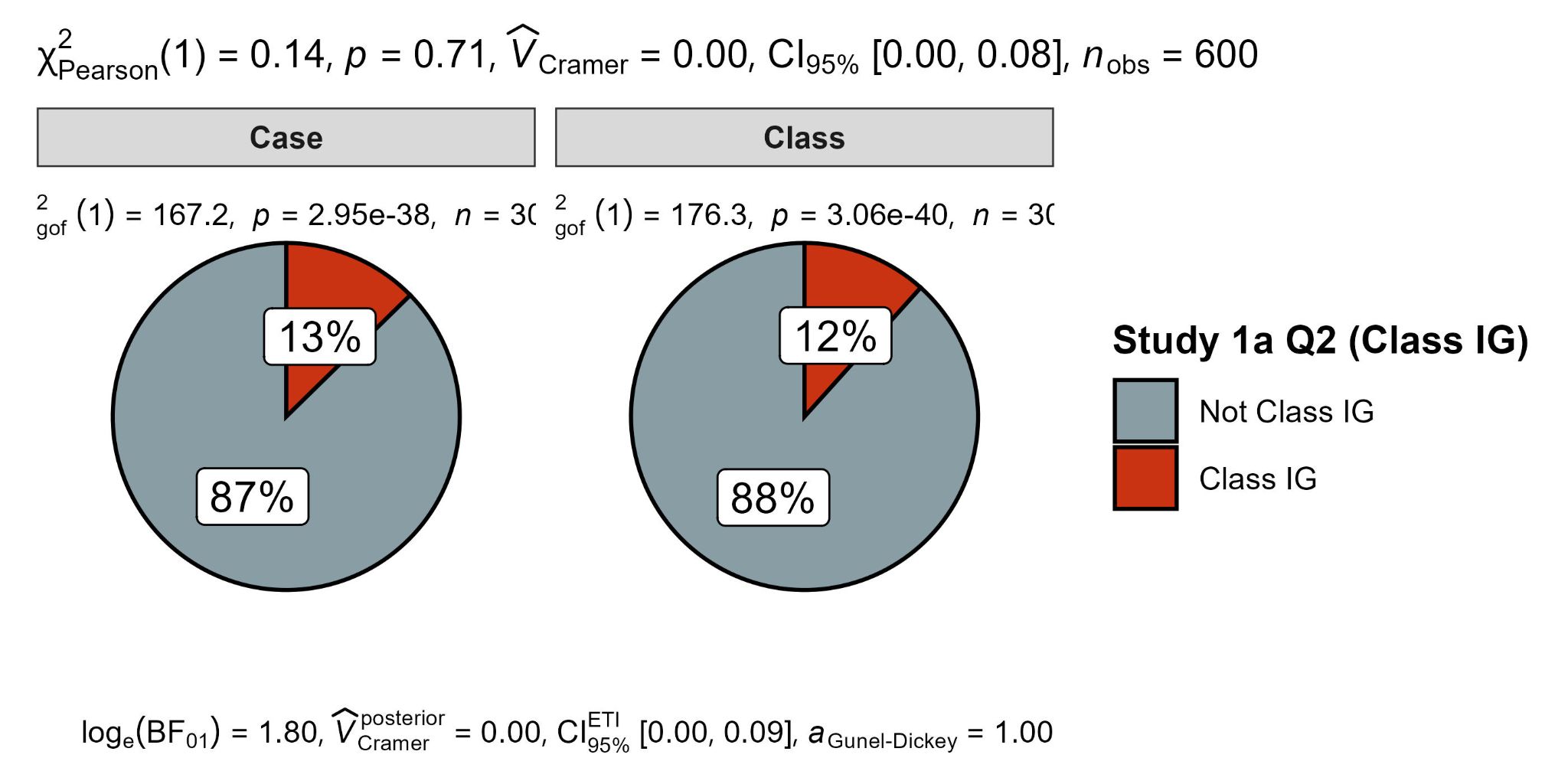
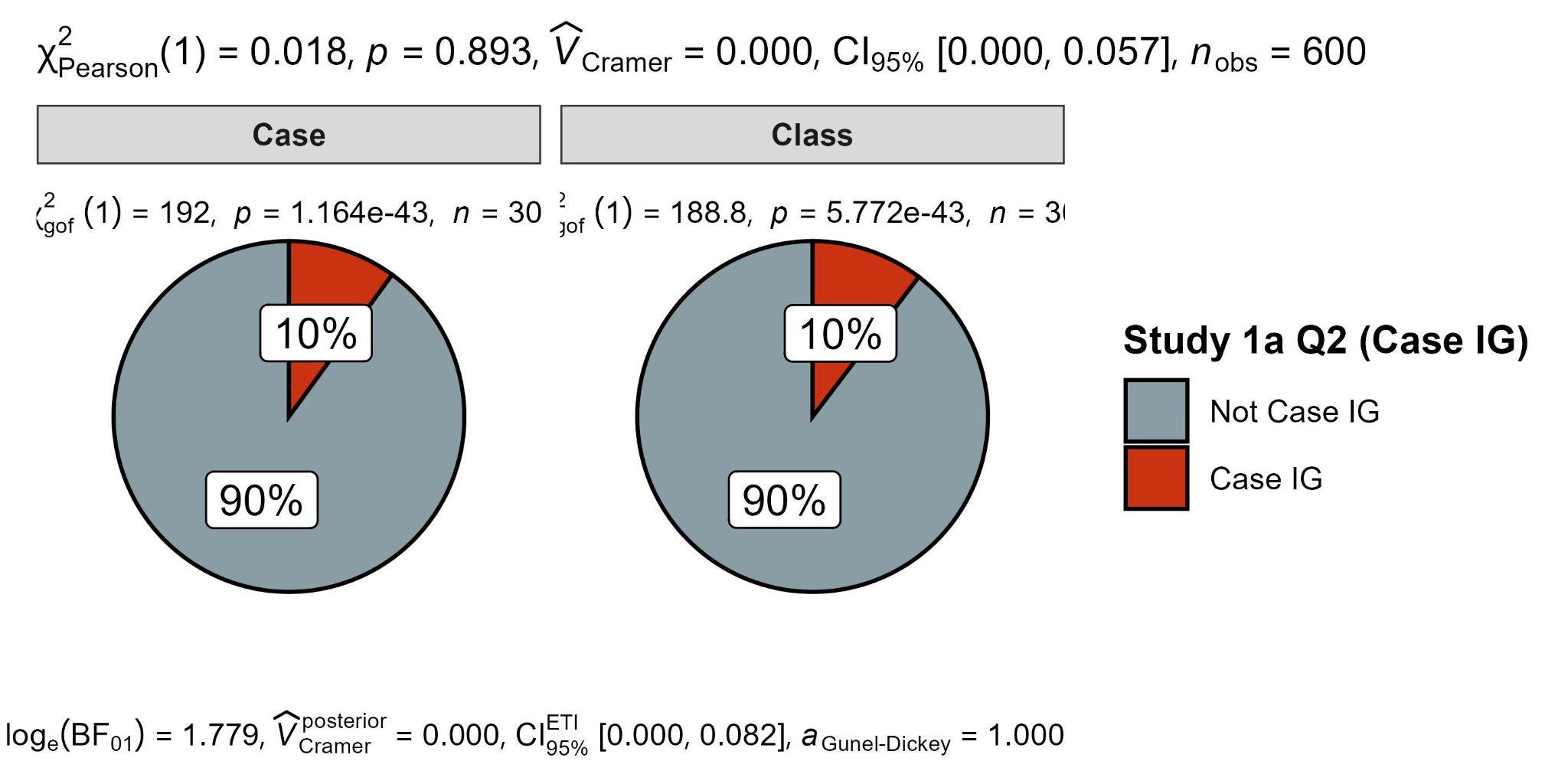


###### Figure 2 *Studies 1a, 1b, 3, and 4: Chi-square tests for ignorance priors (1a/b and 3) and decision (4)*

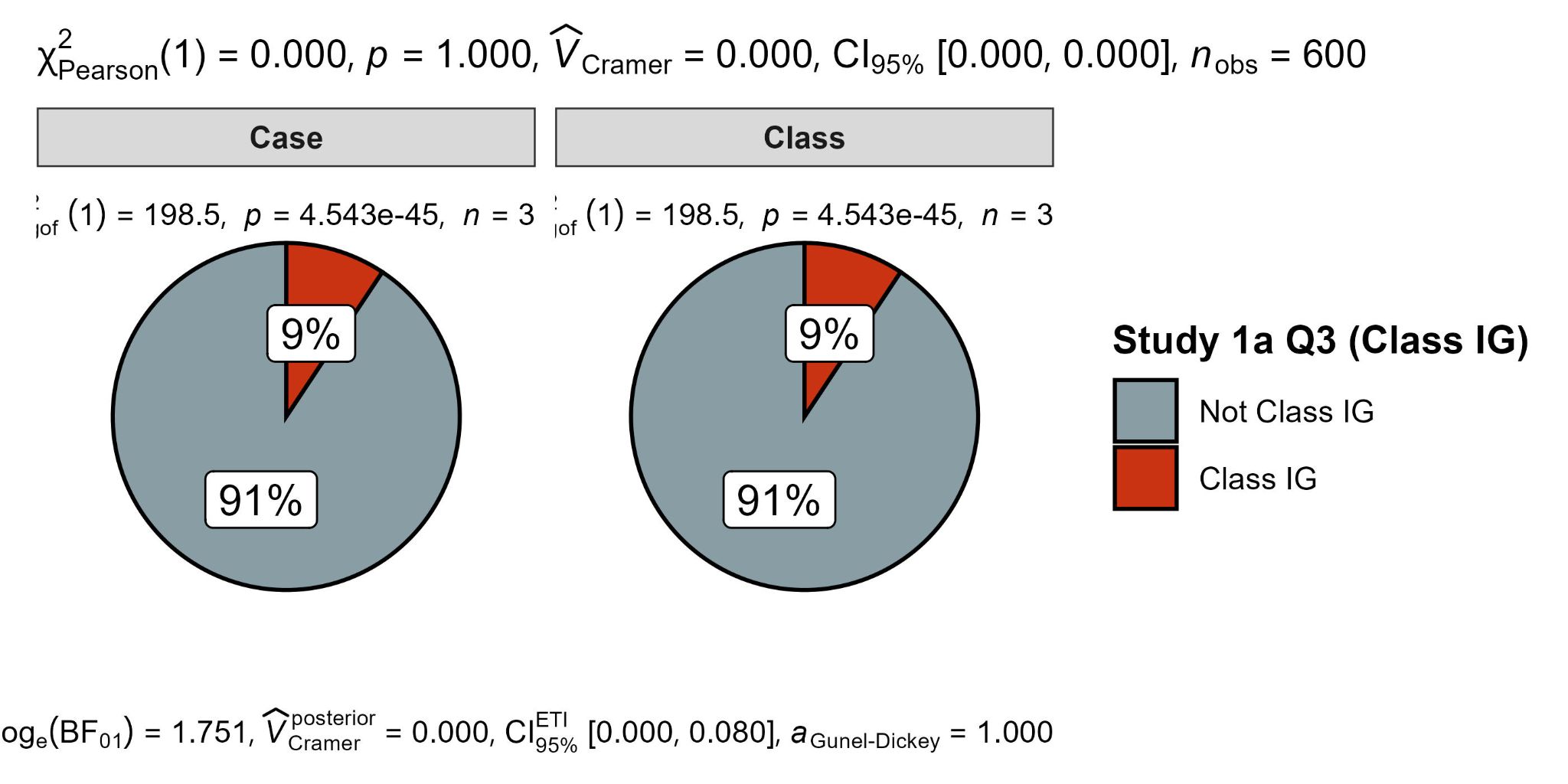
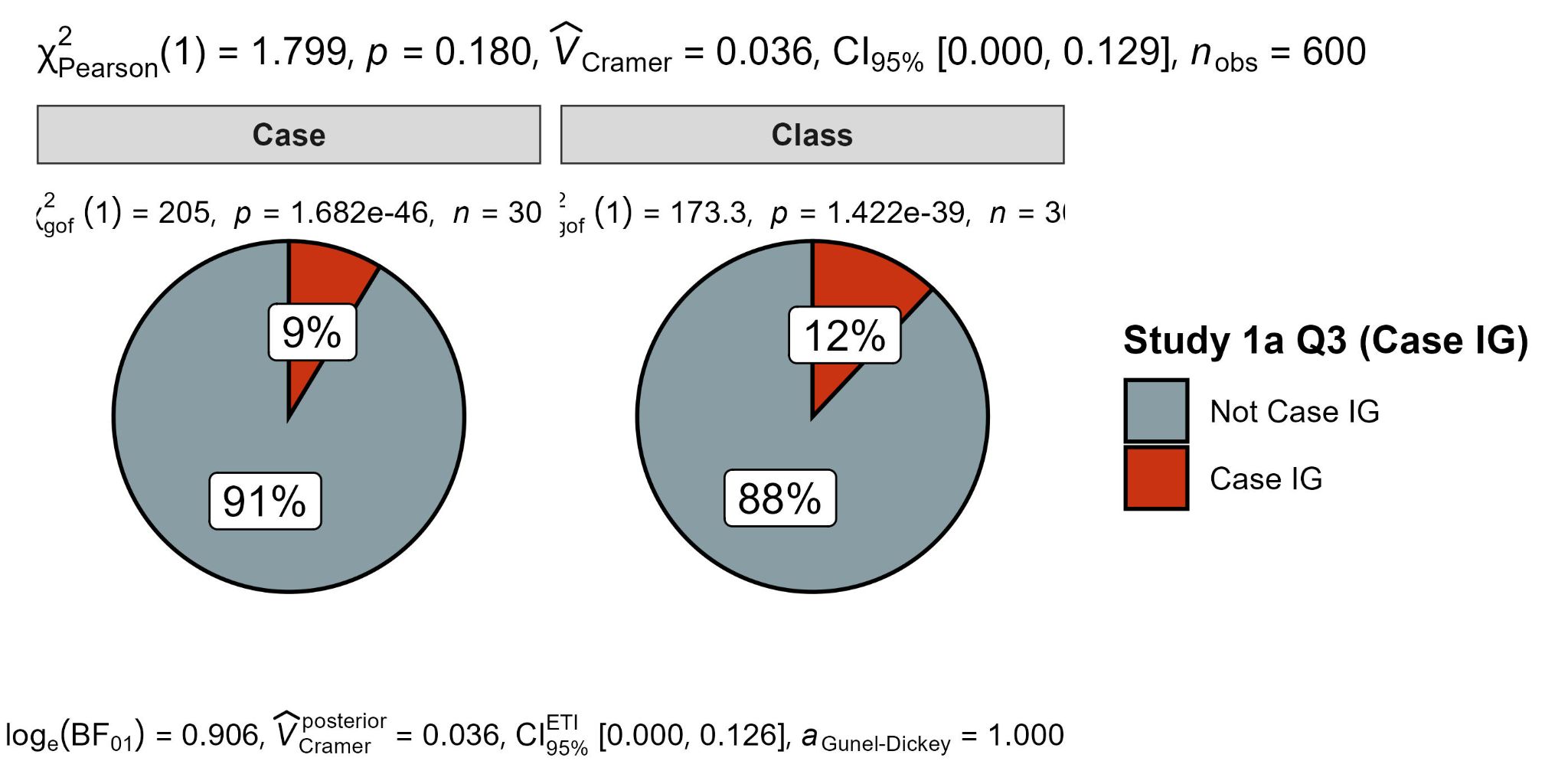
Study 1a Q1



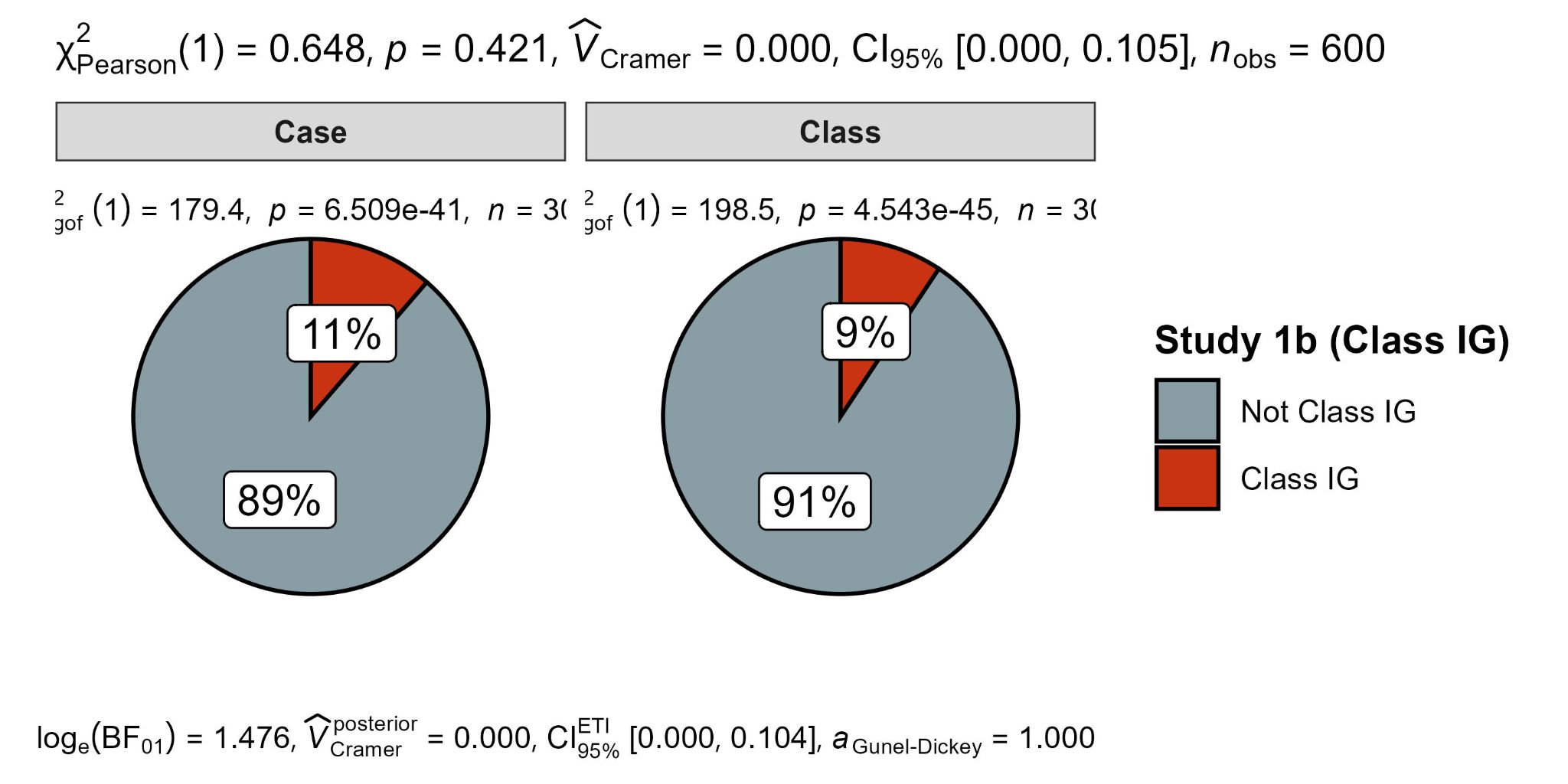
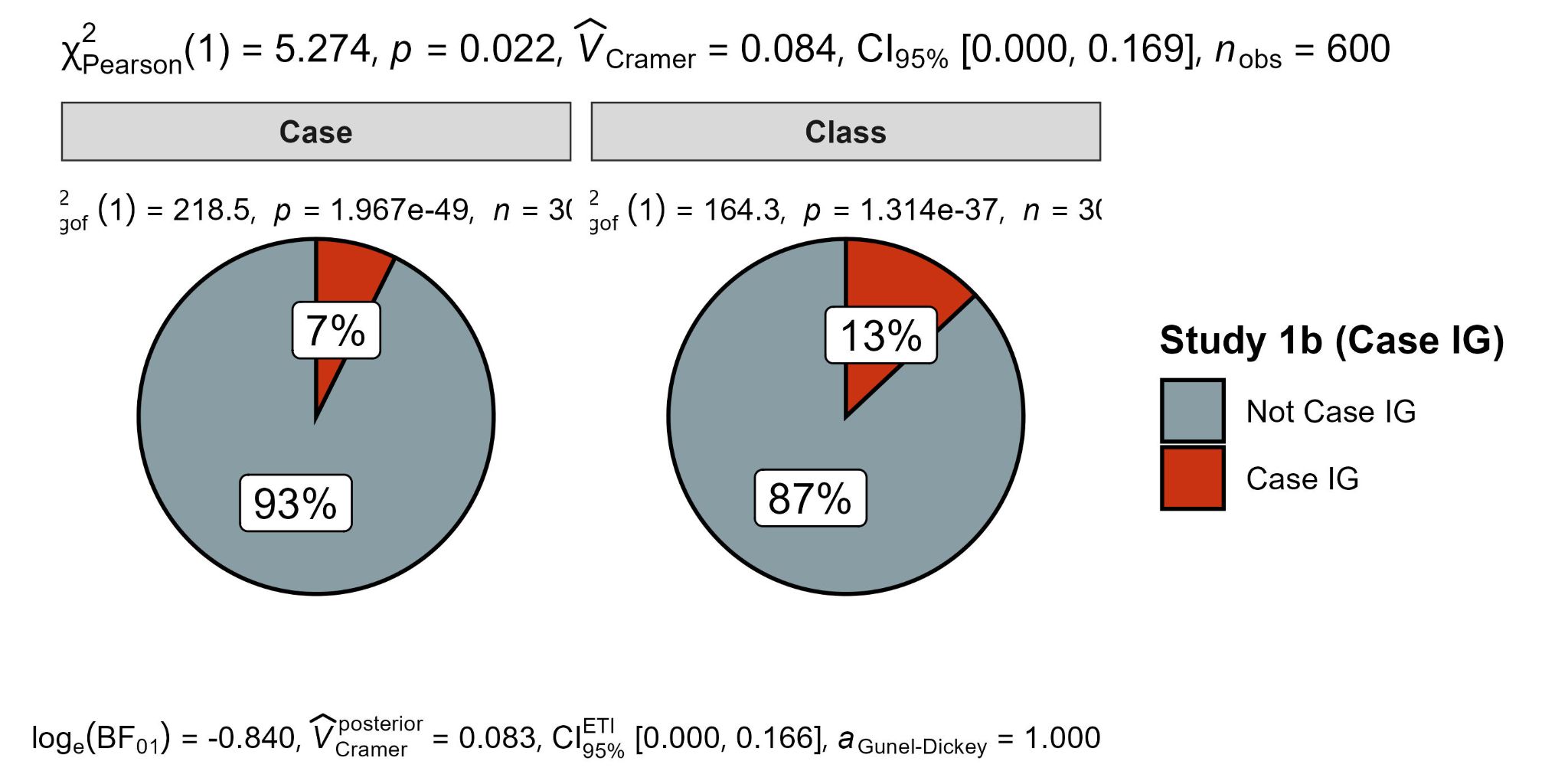
Study 1a Q2



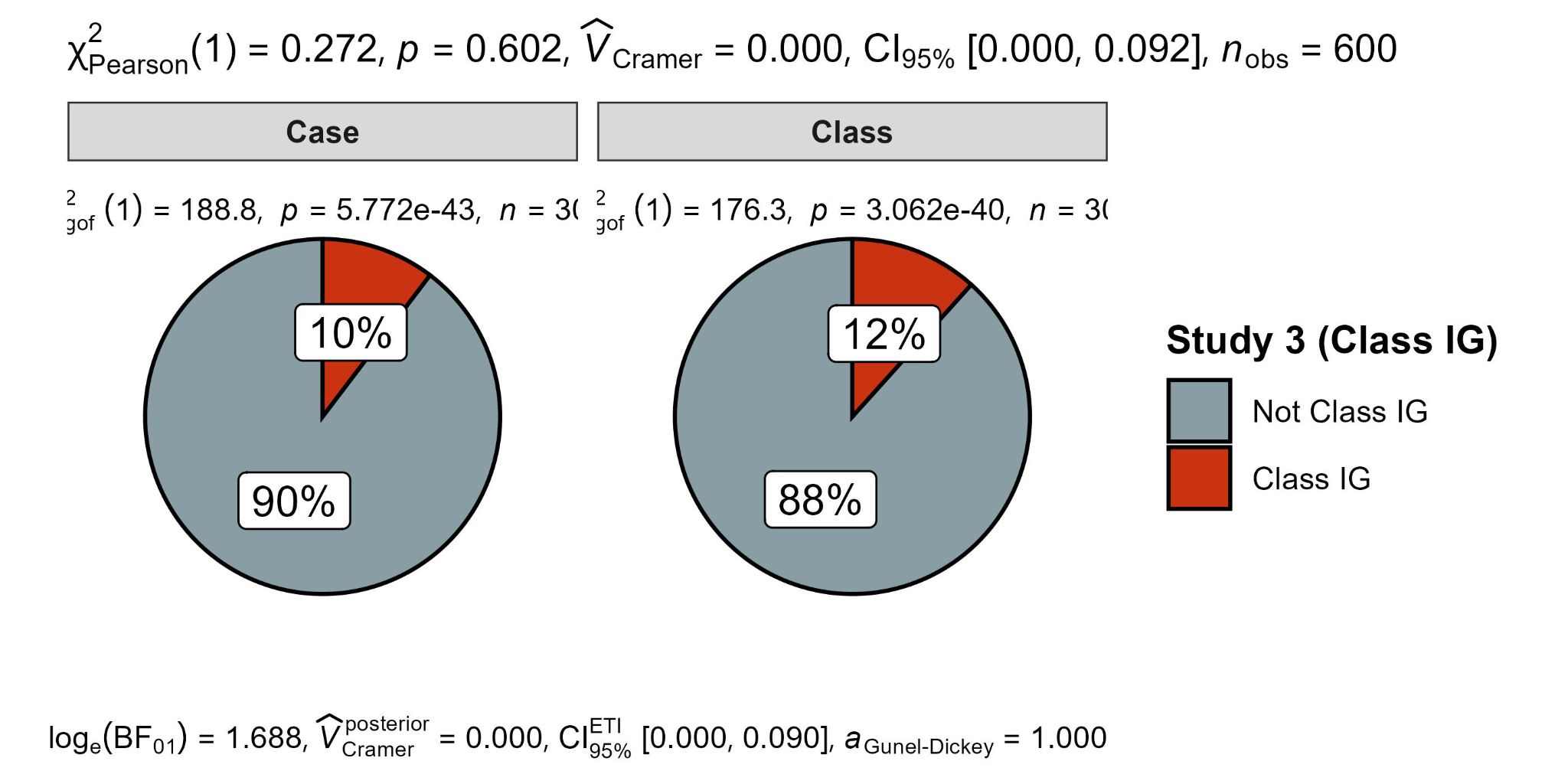
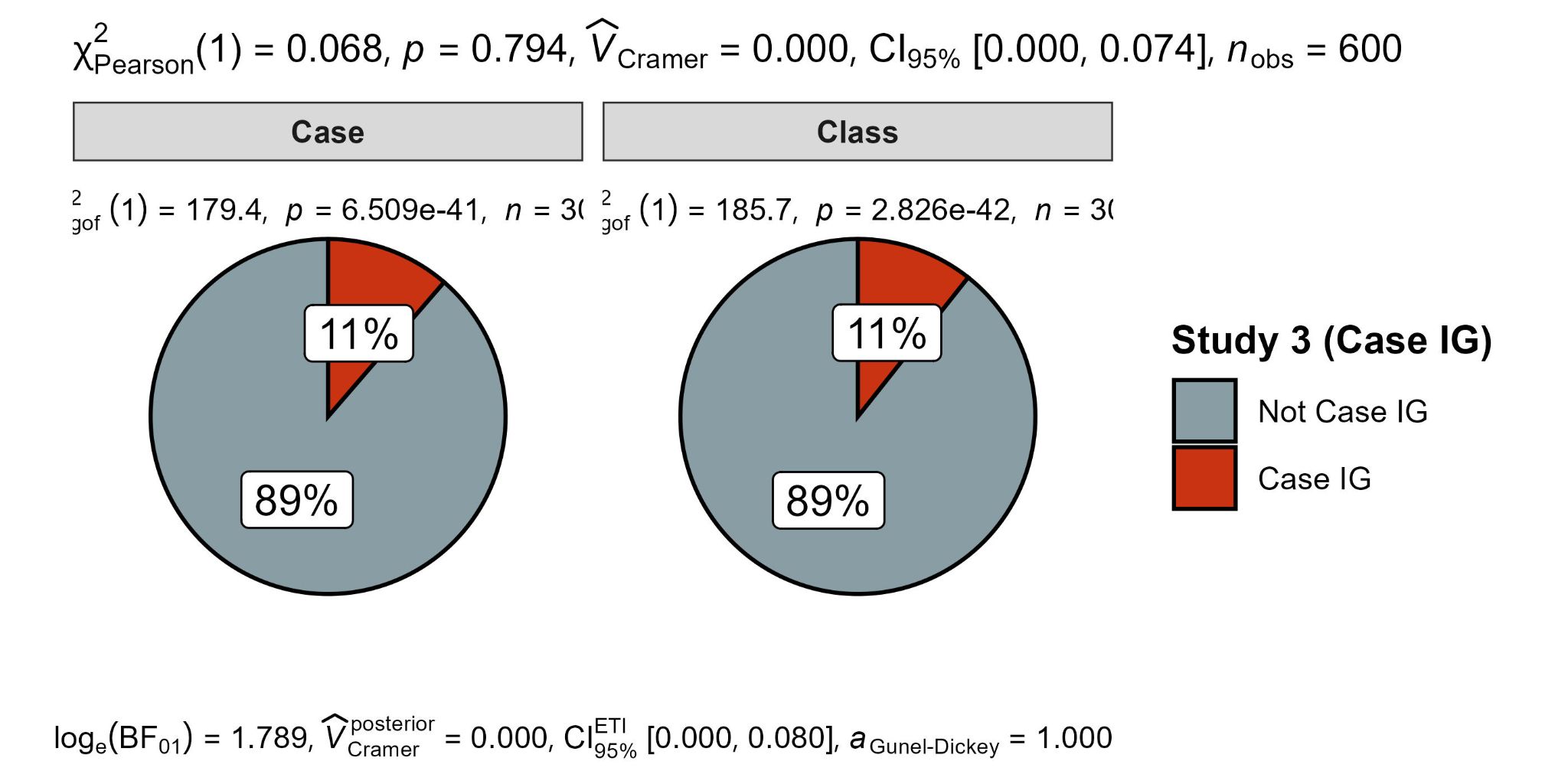
Study 1a Q3



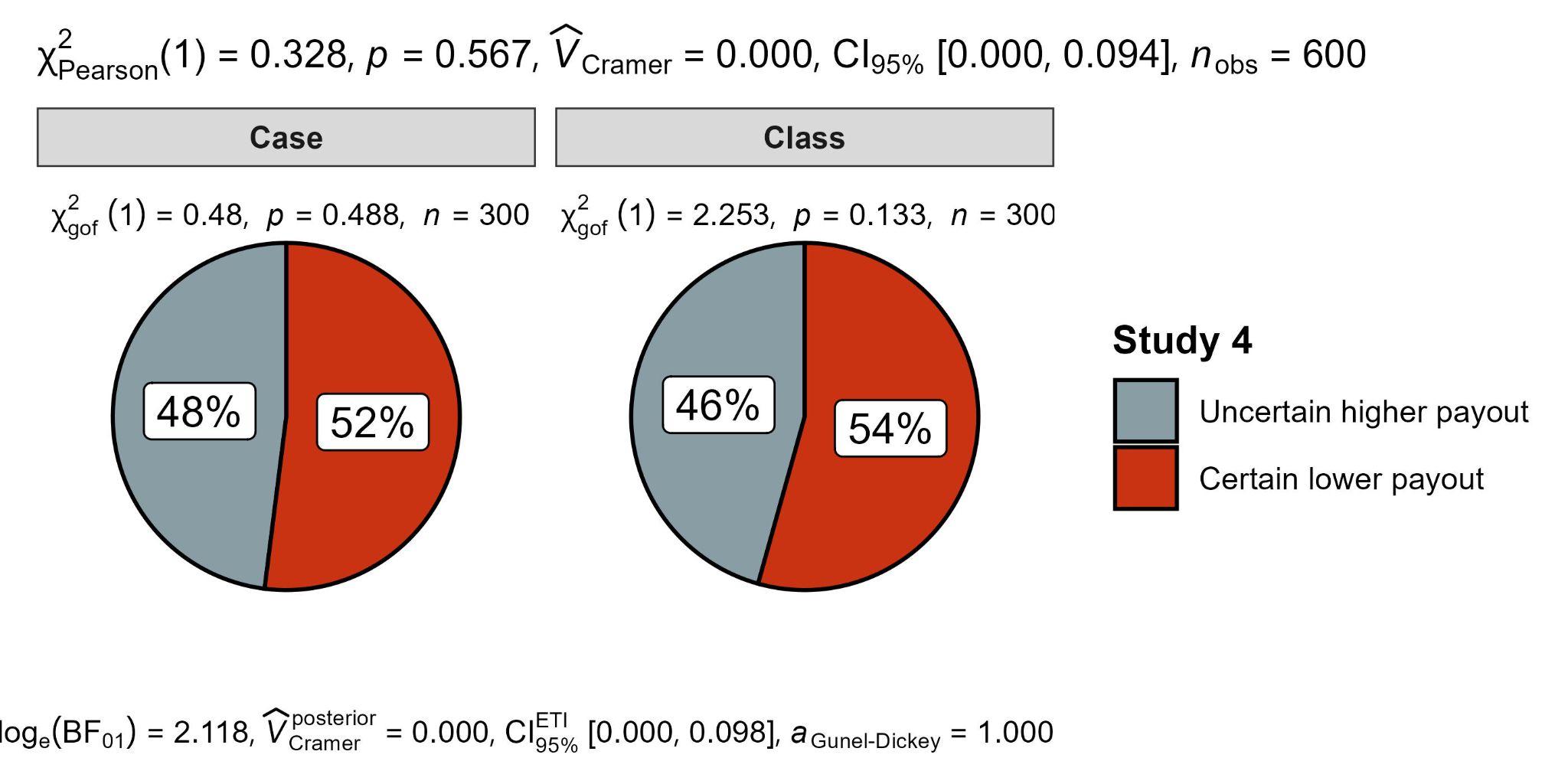
Study 1b



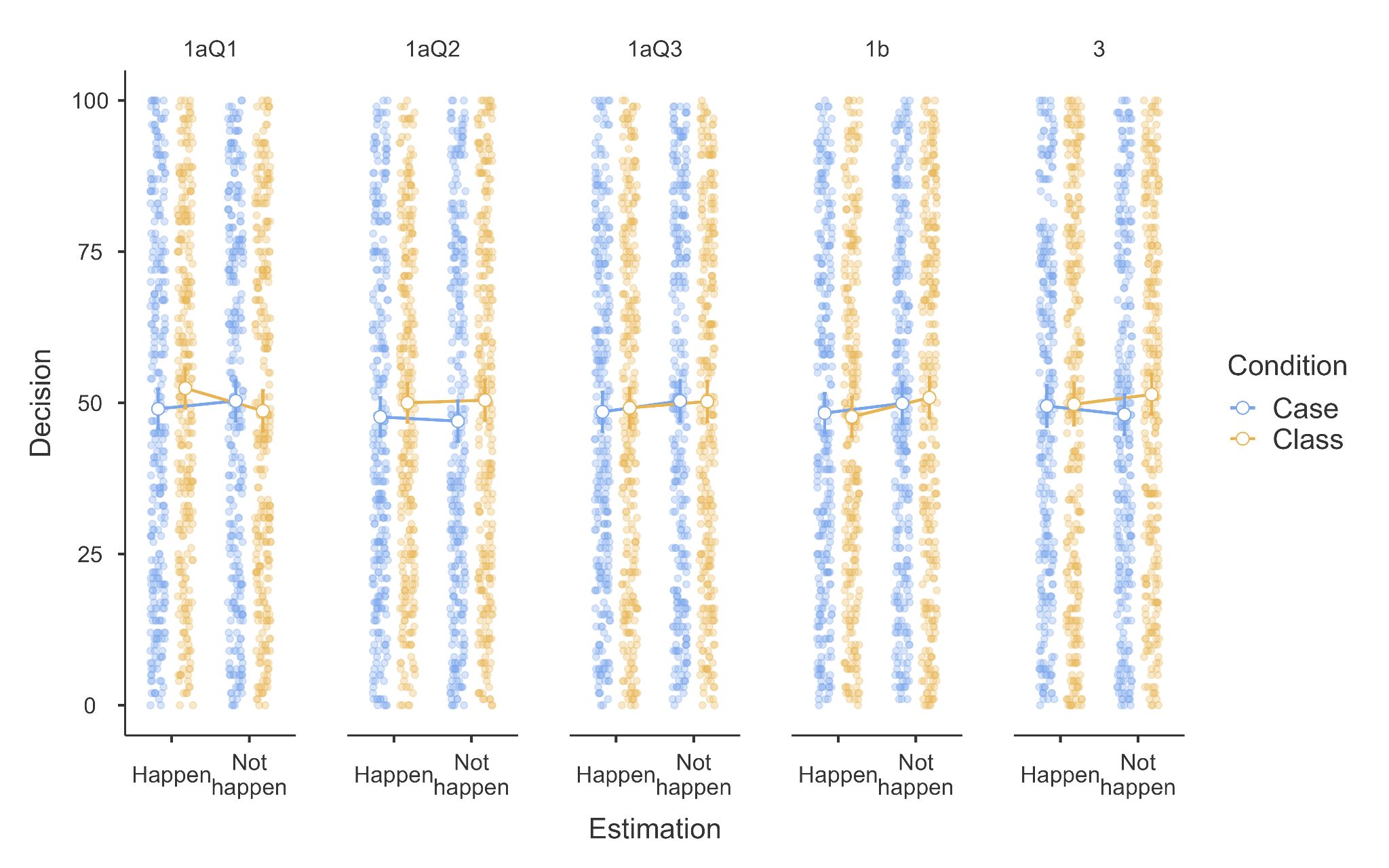
Study 3



Study 4



###### Figure 3 *Three-way mixed ANOVA of condition, scenario, and estimation on probability judgments*



## 

## Replication: Study 4

We conducted a two-proportion *z*-test and found no support for difference in the proportion of participants selecting the second option between the case and the class condition, with proportions of 48% (case) and 46% (class), *z* = - 0.57, *p* = .567, Cohen’s *h* = -0.05, 95% CI [-0.21, 0.11] (see Figure 2).

**Extension: A three-way mixed ANOVA for Studies 1a (cases 1/2/3), 1b, and 3**

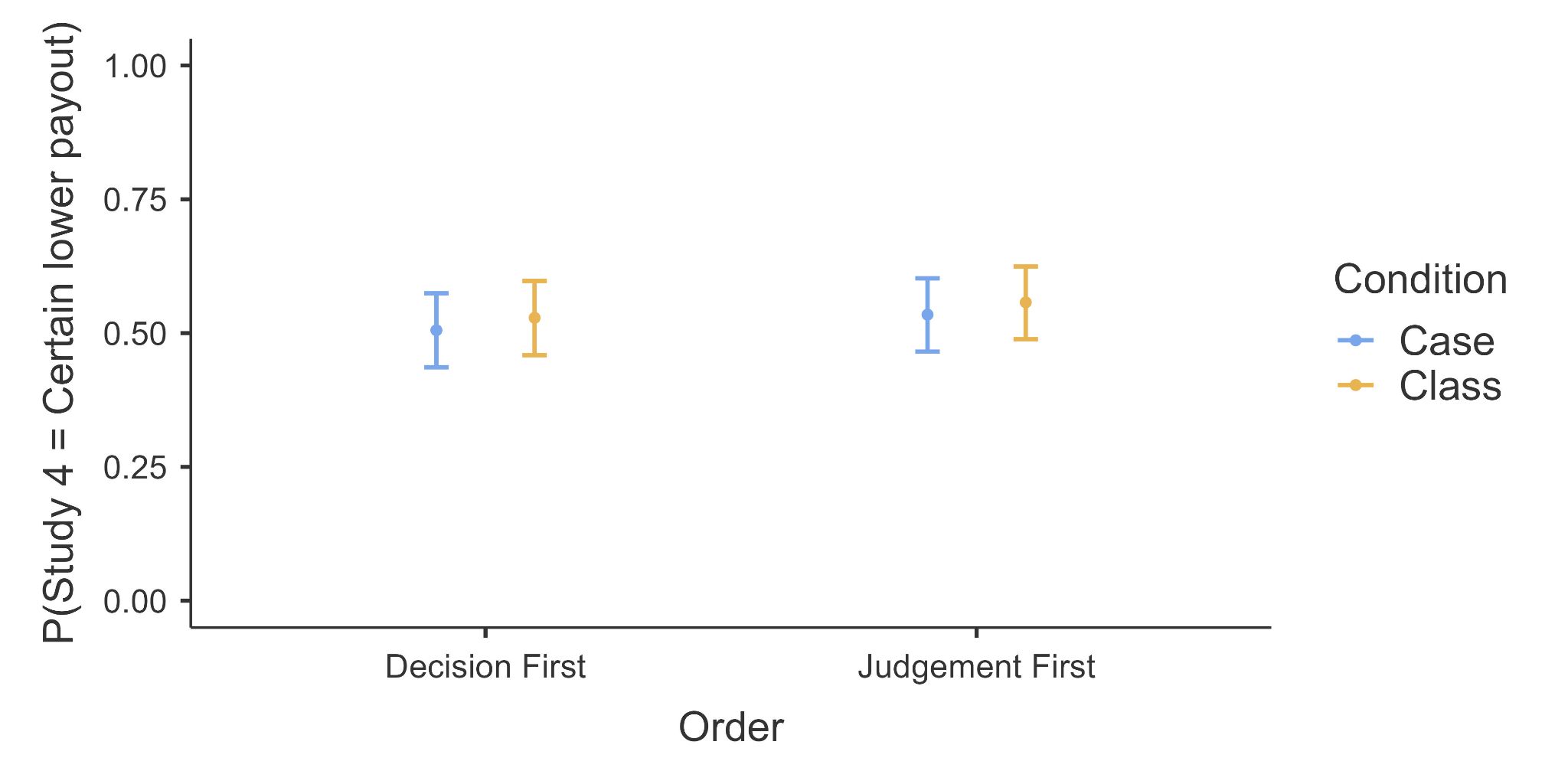
We conducted a three-way mixed ANOVA to examine the effect of condition (case vs. class), question (Study 1a Item 1, Study 1a Item 2, Study 1a Item 3, Study 1b, Study 3), and event type (happen vs. not happen) on participants' probability judgments. The between-subjects factor was condition, and the within-subjects factors were question and event.

We found no support for a main effect of study (*F*(4, 2392) = 0.37, *p* = .831), the interaction between study and condition (𝐹(4, 2392) = 0.48, *p* = .750), the interaction between study, event type, and condition (𝐹(4, 2392) = 0.85, *p* = .494), the main effect of event type (𝐹(1, 598) = 0.48, *p* = .489), the interaction of event type with condition (𝐹(1, 598) = 0.00, *p* = .994, respectively), and the interaction between study and estimation (𝐹(4, 2392) = 0.68, *p* = .604).

**Extension: Order effect**

We conducted a binomial logistic regression to examine the effects of order (Judgement First vs. Decision First) and condition (class vs. case) on the likelihood of participants choosing between two financial options: Option 1 (receive $10 for sure) and Option 2 (receive $50 if the stock whose price per share rises the most on the DJIA today is IBM). We found … [...] (See Figure 4).

###### Figure 4 *Display order analysis: Study 1a-Q3 vs, Study 4*



###### 

###### Table 13 *Replication: Comparison of effects (Cohen’s h and Cohen’s d) between target article and replication.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Replication | Target article |  |
| **Study** |  | **DV** | **Cohen's *h* and CI** | **Cohen's *h* and CI** | **Interpretation** |
| 1a | Item 1 | probability estimation | -0.06 [-0.36, 0.21] | 0.65 [0.37, 0.93] | no-signal - inconsistent |
|  | Item 2 | probability estimation |  |  |  |
|  | Item 3 | probability estimation |  |  |  |
| 1b |  | probability estimation | -0.07 [-0.36, 0.21] | 0.67 [0.39, 0.95] | no-signal - inconsistent |
| 3 |  | probability estimation | 0.06 [-0.21, 0.34] | 0.62 [0.34, 0.90] | no-signal - inconsistent |
| 4 |  | decision outcome | -0.04 [-0.32, 0.24] | -0.32 [-0.60, -0.05] | no-signal - consistent |
| **Study** |  | **DV** | **Cohen's *d* and CI** | **Cohen's *d* and CI** | **Interpretation** |
| 1a | Item 1 | probability estimation | -0.12 [-0.28, 0.04] | 0.55 [0.13, 0.96] | no-signal - inconsistent |
|  | Item 2 | -0.08 [-0.24, 0.08] | 0.55 [0.13, 0.96] | no-signal - inconsistent |
|  | Item 3 | -0.02 [-0.18, 0.14] | 0.55 [0.13, 0.96] | no-signal - inconsistent |
| 1b |  | probability estimation | 0.02 [-0.14, 0.18] | 0.87 [0.22, 1.51] | no-signal - inconsistent |
| 3 |  | probability estimation | -0.01[-0.17, 0.15] | 0.47 [0.00, 0.93] | no-signal - inconsistent |

*Note*. CI = 95% confidence intervals. The interpretation of outcome was based on LeBel et al. (2019). The upper panel presents Cohen’s *h* comparing proportions of class-primed responses in the two experimental conditions (Studies 1a, 1b, and 3, H1 “happen” events) and proportions of choice option (Study 4). The lower panel presents Cohen’s *d* comparing probability judgements (Studies 1a, 1b, and 3)

## Replication Evaluation

To assess the reproducibility of our results compared to those of the target article, we employed the paradigm outlined by LeBel et al. (2019), examining the existence of a signal and comparing confidence intervals with the effect size reported in the target article. We summarized the outcomes of this comparison in Table 13. Following the framework proposed by LeBel et al. (2019), Fox and Rottenstreich (2003) detected signals for Cohen’s *h* in all four studies we replicated, as well as Cohen’s *d* in Studies 1a, 1b, and 3. In our replication, no signal was detected for Cohen’s *d*, and the effect sizes were inconsistent with those in the target article for Studies 1a, 1b and 3. For Cohen’s *h*, no signal was detected across all studies, and the effect sizes were inconsistent with those reported in the original article for Studies 1a, 1b, and 3, but consistent in Study 4. We inferred that our replication failed to reproduce any of the findings outlined in the target article, signaling an unsuccessful replication.

## Extensions: Studies 1a, 1b, and 3 (exploratory analysis)

We examined whether participants' likelihood judgments of negative complementary events (H2) displayed a consistent or divergent pattern of reliance toward partition-dependent priors across Studies 1a, 1b, and 3. [...]

###### Table 14 *Study 1a, b, and 3: results of Mann-Whitney U and Welch’s t test comparing probability estimation of H2 (negative) target events*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Study | | 1a | | | 1b | 3 |
|  |  | Item 1 | Item 2 | Item 3 |
| Mann-Whitney *U* | *U* | 43563.00 | 41976.00 | 44889.00 | 44138.00 | 41980.00 |
|  | *p* | .499 | .154 | .958 | .685 | .155 |
| Welch’s *t* | *t* | 0.68 | -1.02 | 0.04 | 0.27 | -1.49 |
|  | *df* | 597.65 | 597.65 | 596.85 | 596.15 | 597.90 |
|  | *p* | .497 | .152 | .968 | .694 | .161 |
|  | Cohen’*s d* and 95% CI | 0.06  [-0.10, 0.22] | -0.12  [-0.28, 0.04] | 0.00  [0.00, 0.00] | -0.03  [-0.19, 0.13] | -0.11  [-0.27, 0.05] |

*Note.* Probability estimations were recorded in percentages.

# Discussion

[Discussion will be completed in Stage 2 following data collection]

# Conclusion

[Conclusion will be completed in Stage 2 following data collection]

# References

Ben-Shachar, M., Lüdecke, D., Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software, 5*(56), 2815. doi:10.21105/joss.02815

Bulus, M. (2023). Pwrss: Statistical power and sample size calculation tools. *R Package Version 0.3*, *1*.

Coles, N. A., Tiokhin, L., Scheel, A. M., Isager, P. M., & Lakens, D. (2018). [The costs and benefits of replication studies](https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/costs-and-benefits-of-replication-studies/65E3C8C7DAF24F06FE58E800A4F4DA69). *Behavioral and Brain Sciences*, *41*. <https://doi.org/10.1017/S0140525X18000596>

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>

Feldman, G. (2023). Registered Report Stage 1 manuscript template. <https://doi.org/10.17605/OSF.IO/YQXTP>

Feng, Z., Liu, Y., Wang, Z., & Savani, K. (2020a). Let’s choose one of each: Using the partition dependence effect to increase diversity in organizations. *Organizational Behavior and Human Decision Processes*, *158*, 11–26. <https://doi.org/10.1016/j.obhdp.2020.01.011>

Fox, C. R., Ratner, R. K., & Lieb, D. S. (2005). How Subjective Grouping of Options Influences Choice and Allocation: Diversification Bias and the Phenomenon of Partition Dependence. *Journal of Experimental Psychology: General*, *134*(4), 538–551. <https://doi.org/10.1037/0096-3445.134.4.538>

Fox, C. R., & Rottenstreich, Y. (2003). Partition Priming in Judgment Under Uncertainty. *Psychological Science*, *14*(3), 195–200. [https://doi.org/10.1111/1467-9280.0243](https://doi.org/10.1111/1467-9280.02431)1

Fox, C. R., & Clemen, R. T. (2005). Subjective Probability Assessment in Decision Analysis: Partition Dependence and Bias Toward the Ignorance Prior. *Management Science*, *51*(9), 1417–1432. <https://doi.org/10.1287/mnsc.1050.0409>

Jané, M., Xiao, Q., Yeung, S., Ben-Shachar, M. S., Caldwell, A., Cousineau, D., Dunleavy, D. J., Elsherif, M., Johnson, B., Moreau, D., Riesthuis, P., Röseler, L., Steele, J., Vieira, F., Zloteanu, M., & Feldman, G. (2024). Guide to Effect Sizes and Confidence Intervals. http://dx.doi.org/10.17605/OSF.IO/D8C4G

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, *106*(1), 62–88. <https://doi.org/10.1037/0033-295x.106.1.62>

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*, 389-402. <https://doi.org/10.1177/2515245918787489>

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, 3, 1-9. DOI: 10.15626/MP.2018.843 <https://doi.org/10.15626/MP.2018.843>

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* Cambridge University Press. [https://doi.org/10.1017/CBO9781139087759](https://psycnet.apa.org/doi/10.1017/CBO9781139087759)

Li, M., & Feldman, G. (2024). Revisiting diversification bias and partition dependence: Replication and extensions of Fox, Ratner, and Lieb (2005) Studies 1, 2, and 5. Registered Report Stage 1 in-principle acceptance from Peer Community in Registered Reports. <https://doi.org/10.17605/OSF.IO/FUJSV>

Lüdecke, D. (2019). \_esc: Effect Size Computation for Meta Analysis (Version 0.5.1). <https://doi.org/10.5281/zenodo.1249218>

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., ... & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, 73(1), 719-748. <https://doi.org/10.1146/annurev-psych-020821-114157>

Petrov, N. B., Chan, Y. K. M., Lau, C. N., Kwok, T. H., Chow, L. C. E., Lo, W. Y., Song, W., & Feldman, G. (2023). Sunk Cost Effects for Time Versus Money: Replication and Extensions Registered Report of Soman (2001). *International Review of Social Psychology,* 36(1): 17, 1–18. DOI: https://doi. org/10.5334/irsp.883

R Core Team (2024). R: A Language and environment for statistical computing. (Version 4.3.3) [Computer software]. Retrieved from <https://cran.r-project.org>.

Reichelson, S., Zax, A., Bass, I., Patalano, A. L., & Barth, H. C. (2017). Partition dependence in consumer choice: Perceptual groupings do not reliably shape decisions. *Psychonomic Bulletin & Review*, *25*(3), 1178–1183. <https://doi.org/10.3758/s13423-017-1326-4>

See, K. E., Fox, C. R., & Rottenstreich, Y. S. (2006). Between ignorance and truth: Partition dependence and learning in judgment under uncertainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1385–1402. <https://doi.org/10.1037/0278-7393.32.6.1385>

The JAMOVI project (2023). JAMOVI. (Version 2.4.8) [Computer Software]. Retrieved from <https://www.jamovi.org>.

Vonasch, A. J., Hung, W. Y., Leung, W. Y., Nguyen, A. T. B., Chan, S., Cheng, B. L., & Feldman, G. (2023). ‎“Less Is Better” in Separate Evaluations Versus “More Is Better” in Joint Evaluations: Mostly ‎Successful Close Replication and Extension of Hsee (1998)‎. *Collabra: Psychology*, 9(1). doi: 10.1525/collabra.77859 <https://doi.org/10.1525/collabra.77859>

Yeung, S. K., & Feldman, G. (2022). Revisiting the temporal pattern of regret in action versus inaction: Replication of Gilovich and Medvec (1994) with extensions examining responsibility. *Collabra: Psychology*, *8*(1), 37122. <https://doi.org/10.1525/collabra.37122>

Zhu, M., & Feldman, G. (2023). Revisiting the Links Between Numeracy and Decision Making: Replication Registered Report of Peters et al.(2006) With an Extension Examining Confidence. *Collabra: Psychology*, *9*(1). <https://doi.org/10.1525/collabra.77608>

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *The Behavioral and Brain Sciences*, *41*, 1-e120. <https://doi.org/10.1017/S0140525X17001972>