**Lure of choice revisited:   
Replication and extensions Registered Report of Bown et al. (2003)  
[Stage 1]**

Angela Nga Yi Chan  
ORCID: 0009-0004-4331-0460  
Department of Psychology, University of Hong Kong, Hong Kong SAR  
[u3578165@connect.hku.hk](mailto:u3578165@connect.hku.hk) ; [ngayic@gmail.com](mailto:ngayic@gmail.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:

Angela Nga Yi Chan was a thesis student at the University of Hong Kong during the academic year 2023-4.

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on judgement 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:

Angela Nga Yi Chan conducted the replication 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 licence is applied to the AAM arising from this submission. ([clarification](https://bit.ly/rrs-primer))

## Citation of the target article for replication:

Bown, N. J., Read, D., & Summers, B. (2003). The lure of choice. *Journal of Behavioral Decision Making*, 16(4), 297–308. doi: 10.1002/bdm.447

## Acknowledgements (with ORCID)

## Contributor Roles Taxonomy

|  |  |  |
| --- | --- | --- |
| **Role** | **Angela Nga Yi Chan** | **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 |
| Visualisation | X |  |
| Writing-original draft | X |  |
| Writing-review and editing |  | X |

# Stage 1 Snapshot

**‎Provisional title.**

Lure of choice revisited: Replication and extensions Registered Report of Bown et al. (2003)  
**Authors and affiliations.**

Nga Yi (Angela) Chan, Gilad Feldman; Department of Psychology, University of Hong Kong

**Research question(s) and/or theory.**

We aim to conduct a replication and extension of Bown et al. (2003) Studies 1, 2, and 3 and will test their theory and hypotheses. The studies showed that people prefer options that allow them to make further choices over those that do not, even when the extra choices cannot improve the ultimate outcome.

Research questions: whether an option with more choices lures towards that option (Study 1); whether pairing an item with a lure choice in a larger choice set affects the ultimate choice made by individuals; are the effects of dominated lure and conflicted lure the same (Study 2); whether the option to ‘choose to choose’, without specifying a particular option, result in higher likelihood of decision change compared to situations that require immediate choice (Study 3).

**Hypotheses.**

Our replication of Bown et al. (2003) Studies 1, 2, and 3 will follow their hypotheses.

Studies 1 and 2: An option is more likely to be chosen when grouped in an option pair than when option is presented on its own (Study 2: for both dominated and conflicted options).

Study 3: In a Monty-Hall scenario, participants are more likely to switch when presented with an option pair than when only presented with singular options.

**Study design and methods.**

We followed the target article’s design across Studies 1, 2, and 3, in a unified design combining the three experiments in a single data collection.

Study 1 has two scenarios: nightclub, and bank. Three options are organised into two choice sets: 1 option and 2 options. Options are selected more frequently when presented as part of a 2-option set (more choice).

Study 2 casino gambling scenario: Similar to Study 1, betting options are presented in tables, one with choice (2 options) and another without (1 option). Add to Study 1 in that lure option either dominates or is conflicted with reasonable options.

Study 3 Monty Hall problem: Three conditions: Traditional three-door problem (3 door), four-door with a lure of more choice (4 CAC), 4-door without choice lure (4 CAD). Participants first choose a door, one of the remaining doors that does not contain the prize is opened.

We will recruit participants from CloudResearch/Prolific using best practices to ensure high quality data. We initially determine required sample size with a power analysis of the originally reported effects (95%, 0.05) employing Simonsohn (2015) and 10% for exclusions.

**Key analyses that will test the hypotheses and/or answer the research question(s).**

We followed the target article’s data analyses (e.g., chi-square test; Fisher’s exact test).

In Study 1 and 2, we compare an option selected when part of a 2-option set compared to when alone.

In Study 3, we compare a bet selected when part of a 2-option set (4 door CAC) compared to when alone in 3 or 4 door (CAD) conditions.

**‎Conclusions that will be drawn given different results.**

We will evaluate the replicability of our findings against the original’s using the LeBel et al. (2019) paradigm (examining signal and comparisons of effect confidence intervals).

**Key references.**

Bown et al. (2003). The lure of choice. <https://doi.org/10.1002/bdm.447>

LeBel et al. (2019).<https://doi.org/10.15626/MP.2018.843>

# PCIRR-Study Design Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Question | Hypothesis | Sampling plan | Analysis plan | Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis | Interpretation given different outcomes | Theory that could be shown wrong by the outcomes |
| Are (and to what extent) are people attracted to options that allow them to make further choices? | An option is more likely to be chosen when it is presented as part of a larger choice providing decision-makers with more choice. | Based on a power analysis, aiming for 1026 participants, well-powered enough to detect *w* = 0.16 for a six-conditions chi-squared goodness-of-fit test for our extension (95% power, alpha = 5%, two-tail). See Power analysis section. | Fisher’s exact test,  Chi-square test of independence (Extension) | We conducted a power analysis of the target’s reported effects, with adjustments based on the small-telescopes (Simonsohn, 2015) approach, multiplying these by 2.5 and accounting for possible exclusions. | We examine the replicability of Bown et al. (2003) based on the replication comparison criteria by Lebel et al. (2019). | That the lure of choice influences people’s decision-making. |

# Abstract

[IMPORTANT: Abstract, method, and results were written using a randomised 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.]

The lure of choice is a heuristic that suggests people prefer options that allow them to make further choices over those that do not, even when the extra choices cannot improve the ultimate outcome. In a Registered Report, we conducted a replication and extensions of Studies 1, 2, and 3 from Bown et al. (2003), with a British Prolific sample (*N* = 1000). [The following is a demo placeholder and will be updated following data collection.] We [found/did not find] support for the lure of choice in the nightclub scenario (*V* = X.XX, 95% CI [X.XX, X.XX]), bank scenario (*V* = X.XX, 95% CI [X.XX, X.XX]), casino dominated lure condition (*V* = X.XX, 95% CI [X.XX, X.XX]), conflicted lure condition (*V* = X.XX, 95% CI [X.XX, X.XX]), and the Monty Hall problem scenario (*V* = X.XX, 95% CI [X.XX, X.XX]). Extending the replication, we added neutral control conditions and [found/did not find support] for the neutral control conditions as different from the experimental conditions [...details…]. Overall, we conclude that …. Materials, data, and code are available on: <https://osf.io/e47jh/>

*Keywords:* Lure of choice; heuristic; judgement and decision making; registered report; replication; sequential choice; choice

# Lure of choice revisited: Replication and extensions Registered Report of Bown et al. (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

Bown et al. (2003) demonstrated the “lure of choice” phenomenon, by showing that people prefer options that allow them more choice over those that do not, even when the additional choices cannot improve outcomes.

We conducted a close replication and extension Registered Report of Bown et al. (2003) with two main goals. Our primary goal was to conduct an independent close replication of the lure of choice phenomenon. Our secondary goal was to build on the target’s design and add extensions aiming to improve the target’s methods and gain additional insights.

We begin by introducing the literature on the lure of choice and the chosen article for replication - Bown et al. (2003). We then discuss our motivations for the replication and review the article by Bown et al. (2003) and their theory and hypotheses. Finally, we outline our chosen studies for replication from the target article, the target’s experimental design, and our adaptations and extensions.

## The lure of choice: Bown et al. (2003)

Imagine you are in search of a travel agency to plan your upcoming trip and you have limited time to only visit one agency. You come across two options: Agency A specialises exclusively in one specific type of activity in your chosen destination, an activity which you value, whereas Agency B offers a different activity of similar value to you yet has additional options for other activities, of different or even lower value compared to the main activities by both agencies. Hence, Agencies A and B both have activities you enjoy of similar value, but Agency B has additional options. Which of the two agencies would you prefer?

Bown et al. (2003) argued that in such decisions people would prefer to go for Agency B that has more options, a phenomenon they coined as the “lure of choice”. They demonstrated that people are more likely to choose A when it is paired with another item in a larger choice set ({A or B} and C) than when it is presented in isolation (A and {B or C}). In other words, people prefer options that allow them to make further choices over those that do not.

One explanation for the lure of choice is that it emerges from individuals’ tendency to defer commitment. People tend to overestimate the amount of information required to make decisions (Klein & O’Brien, 2018), and postpone making decisions until they feel they gathered sufficient information (Bastardi & Shafir, 1998). In such circumstances, opting for a path that offers further choices provides them with an opportunity to procrastinate committing to an option, even if the deferral is only temporary.

Another possible explanation is that individuals may be aiming to prioritise decisions that offer more flexibility. The “lure of choice” may be reflective of a heuristic, such as - ‘it is better to keep options open than to burn bridges’. Heuristics may serve a functional purpose, especially in dealing with situations involving uncertainty (Sharfman & Dean, 1997). For example, when one is unsure about a travel destination and the quality of activities it offers, people might perceive it to be better to choose a travel agency that provides more options over one that restricts choices. Thus, people may make decisions based on attributes of the presence or absence of further choices and be lured to make choice-qua-choice decisions. Once they are committed to a choice path, they will start to evaluate the trade-offs among the available options.

However, this heuristic may not always lead to better decisions. In the example of the travel agency, the best experience for the preferred activities might not be by the agency that offers more choices and does a lot of things well enough, but rather by the agency that specialises and does one thing exceptionally well. Therefore, the preferred option is not necessarily found within the larger choice set, and a lack of awareness of the tendency to prefer choice may go counter to one’s own goals, values, and utility. Therefore, the lure of choice is not necessarily better and may end up being worse for the individual than if chosen directly.

### Empirical demonstration

Three of the experiments of Bown et al. (2003) employed the floating lure design, in which participants were presented with two equally desirable targets and a lure. These were paired in such a way that participants had to choose between one target in isolation (targetI) and one target paired with the lure (targetL). Both targets served as targetI and targetL in different conditions. The floating lure design helps separate the lure of choice from the asymmetric dominance effect, also known as the decoy effect or the attraction effect (Huber et al., 1982).

In the nightclub scenario of Study 1, participants were asked to select a nightclub from a choice set consisting of three nightclubs - two targets (Diesel and Cherish) and a “lure” (Atom) separated to two directions (North and South). One direction had two night clubs - a target and the lure, and the opposite direction had the other target nightclub. Therefore, one direction always offered more choice, two clubs compared to one. The three nightclubs differed in terms of entrance fees and levels of enjoyment. The targets differed on price and enjoyment - Club Diesel charged a higher entrance fee but offered a higher level of enjoyment compared to Club Cherish. The “lure”, Club Atom, was dominated by Club Diesel, charging a higher entrance fee but providing the same level of enjoyment. The experiment manipulated which club was coupled with the lure to offer more choice. Participants first chose the direction they wanted to go, and then selected the nightclub they would like to visit if further choices were available. They found that a nightclub target was more likely to be selected when it was coupled with a lure than when presented alone. We summarised the targets and lure options in Figure 1, and the study design and choices in Table 4.

Similarly, in the bank scenario of Study 1, participants were asked to choose from three possible savings options offered by two different banks, meaning that one bank offered more choice - two saving options compared to one. The three accounts differed in terms of their interest rates and notice periods for withdrawal. Specifically, Account B offered a lower interest rate in exchange for an immediate withdrawal, compared to Account A. Unlike the first scenario, in this scenario the “lure”, Account L, was not dominated by either of the targets, but instead served as a compromise between the two targets. The experiment manipulated which bank offered the lure, thereby offering more choice. Participants first chose the bank they wanted to visit, and then selected their preferred saving account if further choices were available. They found that a saving account target was more likely to be selected when it was coupled with a lure in a bank that offered more choice than when in a bank where it was the only option. We summarised the targets and lure options in Figure 2, and the study design and choices in Table 5.

In the casino scenario of Study 2, participants were asked to choose a spinner placed on two different tables. One of the tables had a “lure” spinner, yet Study 2 also manipulated the type of lure - one lure was dominated (worse than both targets), and one lure was conflicted (compromise between the two targets). The spinners differed in terms of their payoffs and probabilities of winning. Target Spinner A had a higher probability of winning but a lower payoff compared to target Spinner B. In one condition, lure Spinner C was a conflicted lure to both target spinners A and B, offering a higher probability of winning but a lower payoff. In the other condition, lure Spinner D was dominated by both spinner targets, featuring a lower probability of winning with the same payoff as Spinner A, and a lower payoff with the same probability of winning as Spinner B. Therefore, this experiment manipulated both which table offered the lure, thereby offering more choice, as well as the type of lure. Participants first chose the table, and then selected the spinner they would like to play if further choices were available. They found that a spinner target was more likely to be selected when it was coupled with a lure in a table that offered more choice than when in a table where it was the only option, regardless of lure type, with a stronger effect for the conflicted lure. We summarised the targets and lure options in Figure 3, and the study design and choices in Table 5.

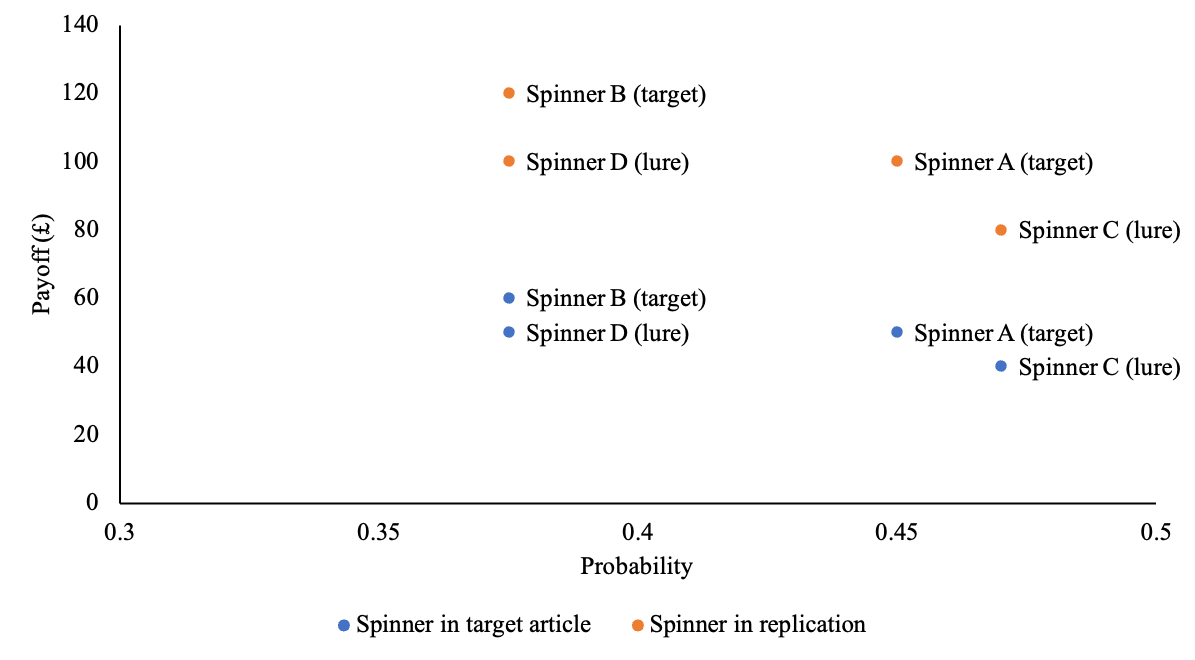
###### Figure 1 *Study 1 Nightclub scenario: Relative dimensions of targets and lure*



*Note*. The figure includes both the values originally used in Bown et al. (2003) and the replication. Prices in the replication were doubled, to adjust for inflation between the years 2003 and 2024.

###### Figure 2 *Study 1 Bank scenario: Relative dimensions of targets and lure*

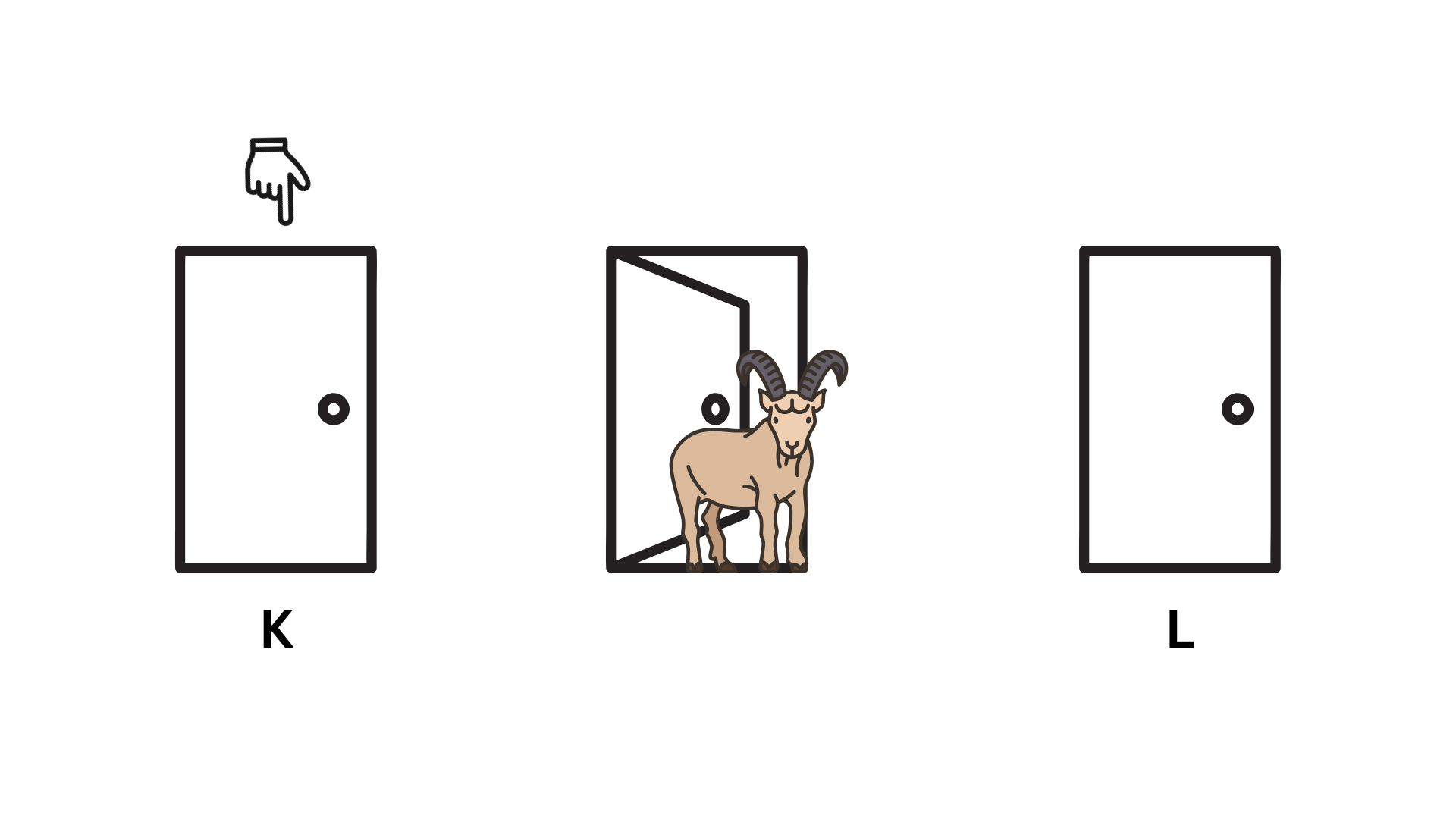
###### Figure 3 *Study 2 Casino scenario: Relative dimensions of targets and lures*

*Note*. The figure includes both the values originally used in Bown et al. (2003) and the replication. Prices in the replication were doubled, to adjust for inflation between the years 2003 and 2024.

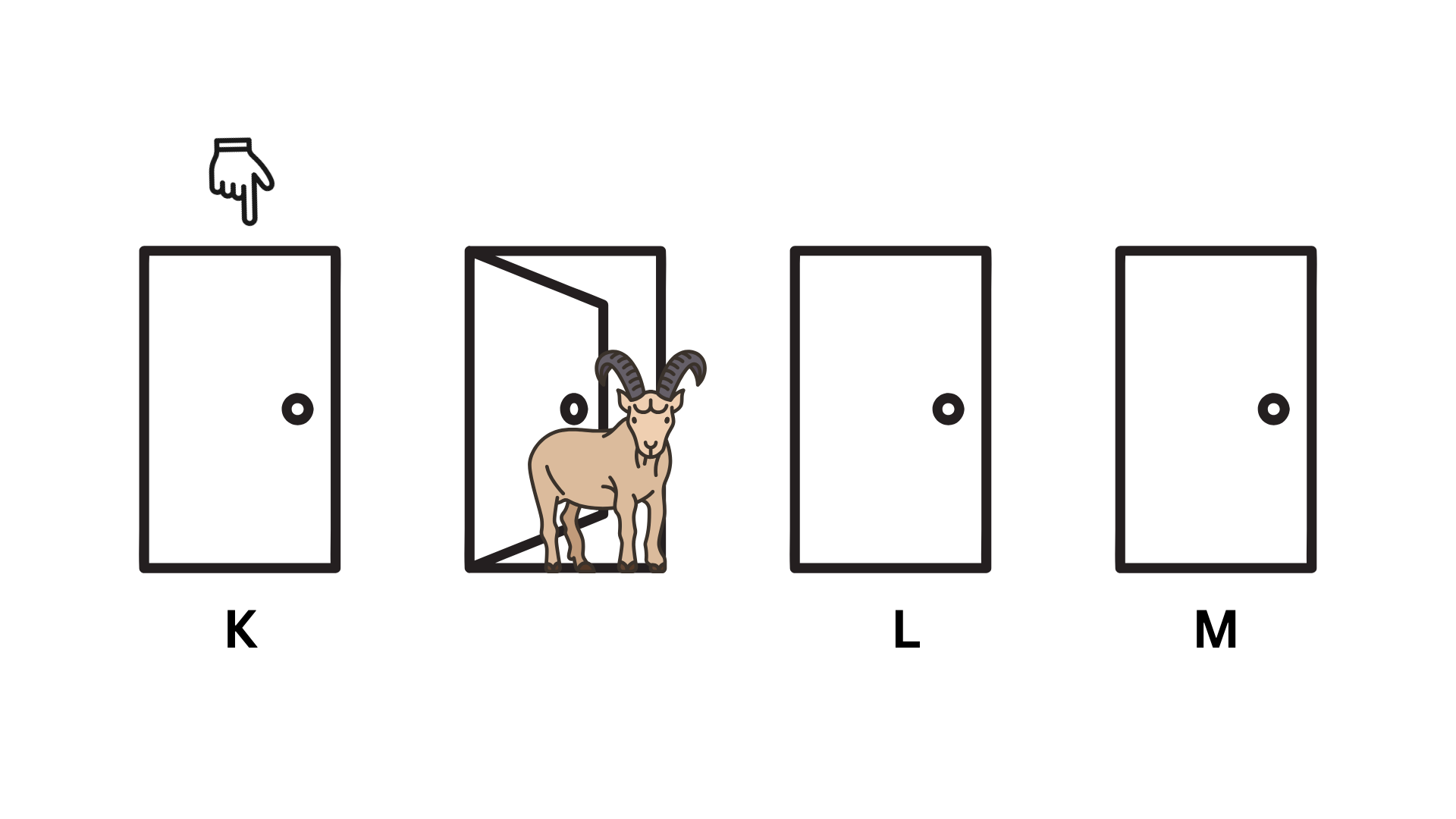
In Study 3, Bown et al. (2003) adjusted the classic Monty Hall problem to investigate the lure of choice. In the classic three-door Monty Hall problem, participants are presented with three doors. The probability of winning by switching doors is 2/3, compared to 1/3 for sticking to the original door. However, most people fail to arrive at an accurate probability of winning and choose to stick with their initial choice instead of switching doors. This phenomenon was evident in the standard 3-door condition, where only 14% of participants switched.

To explore the lure of choice using the Monty Hall problem, Bown et al. (2003) added four-door conditions in which the probability of winning by switching was 3/8, compared to 2/8 for sticking. They showed that the number of participants who switched in the “4-door Choose-a-choice” (CAC) condition, which granted them the opportunity to delay committing to a specific door, was higher than in the “4-door Choose-a-door” (CAD) condition. It demonstrated that retaining the option to choose at a later point by switching in the “4-door CAC” condition overrides the instinct to stick with and commit to the original door in the “4-door CAD” condition. We note that the study found no support for differences in the number of participants who switched (or stuck) between the “3-door” condition and the “4-door CAD” condition. This indicates that the number of options did not affect individuals’ decision-making process, and thus discounted the explanation that the lure of choice stems from the mere summation of options, where individuals select the larger choice set due to the presence of more options. We provided the stimuli used in Study 3 in Figures 4 and 5, and summarised the study design and choices in Table 7.

###### Figure 4 *Study 3 Monty Hall problem: 3-door condition stimulus*



###### Figure 5 *Study 3 Monty Hall problem: 4-door conditions stimulus*



### 

### Lure of choice and asymmetric dominance effect

Bown et al. (2003)’s lure of choice offered a new take on classic phenomena, such as the asymmetric dominance effect.

The asymmetric dominance effect suggests that introducing a third option that is dominated by one target but not dominated by another target enhances the attractiveness and likelihood of selecting the dominating option. For instance, consider a scenario with two cameras, where Camera A is cheaper but of lower quality compared to Camera B. When a third camera, which is more expensive than Camera A but of poorer quality, is introduced to the comparison, Camera A is likely to be chosen more frequently, because it is clearly better than the decoy option, strengthening the reasons for selecting Camera A.

Although both the asymmetric dominance effect and the lure of choice were demonstrated using the presence of a third option, they seem to operate differently. The asymmetric dominance effect revolves around the relative comparison between the options, whereas the lure of choice focuses on the structural arrangement of the options with the larger choice set being more attractive to people. Specifically, in the asymmetric dominance effect the presence of a lure will increase the number of participants choosing the dominating option, regardless of how the options are paired. In contrast, in the “lure of choice” presence of the lure will increase the number of participants choosing targetL, which is different across conditions.

### Lure of choice and rule of regularity

The lure of choice seems to violate the classic rule of regularity, which suggests that the preference for an option should not increase when more options are added to a choice set (Shafir et al., 1993; Tversky & Simonson, 1993). Specifically, the regularity condition states that the frequency of choosing option A from a choice set {A or B or C} should not be higher than the frequency of choosing it from a smaller choice set {A or B}. In Bown et al. (2003), the introduction of option M in the Monty Hall problem resulted in more participants choosing between options L or M, compared to those who chose option L in the standard condition.

The lure of choice suggests that the inclusion of an additional option can substantially increase individuals' preferences for the choice set and deviate from the expected choice patterns observed under regularity.

## Choice of article for replication: Bown et al. (2003)

We embarked on a well-powered close replication and extension Registered Report of Bown et al. (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 Bown et al. (2003) based on several factors: impact on other literature and phenomena, unclear generalizability, the potential for improvements in methodology with extensions by adding missing neutral conditions, its academic and practical impact, and the absence of direct replications thus far.

### Impact and generalizability

The article has had much impact on scholarly research in the areas of judgement and decision-making, and behavioural economics. Above, we noted the links between the lure of choice to asymmetric dominance effect and the violation of the rule of regularity that suggests the need for careful replications. At the time of writing (January 2024), there were 268 citations on Google Scholar, and several follow-up theoretical articles that explored the topic further. For example, Siddiqi (2009) conducted a conceptual replication of the 4-door Monty Hall problem and demonstrated that the lure of choice is not only observed at the individual level but also manifests in market prices.

In addition, the generalizability of “lure of choice” is still unclear. While Bown et al. (2003) demonstrated the lure of choice, even when options are not necessarily better than other available choices, their findings differed from those of earlier studies. For instance, Suzuki (1997, 2000) found that individuals were more inclined to choose an alternative only when all the subsequent options were equal to or better than the no-choice option. In other words, that “lure of choice” seems to take place when the subsequent options are of equal or greater value.

Furthermore, other research found that people preferred isolated items over larger choice sets (Brenner et al. 1999; Sood et al. 2004). When the choices become complex, rendering it difficult for individuals to differentiate between the options, they tend to close off some options and favour a smaller subset of options over a larger array. Also, when individuals encounter lone-versus-group choices, they tend to identify their favourite option within the group choices first, and then compare it with the isolated option. In this case, the group favourite has already been compared to other options within the group in the first step, which highlights all its disadvantages or losses. In contrast, the lone option is only directly compared to the group favourite, which may not expose it to the same extent of loss-emphasising comparisons. This can lead individuals to perceive the group favourite as less attractive compared to the lone option, even though it may have been the preferred choice among the group. Given the divergent findings, it would be best to revisit and validate all studies.

Bown et al. (2003) has also been shown to have practical implications. For instance, Bryant et al. (2007) emphasised that framing healthcare decisions as choices may lead to suboptimal decision-making processes, potentially resulting in regret and choices that are not the best for their patients.

Despite its impact and the theoretical and empirical gaps, to the best of our knowledge, there have been no published independent direct replications of Bown et al. (2003), though not for lack of interest.

### Needed improvements to the studies in the target article

As we analysed the article we realised that some methodology and analyses were in need of revisiting and improvement. Although the target article conducted a pre-test to ensure the two targets in each floating lure design scenario are roughly equally favourable, we thought it would be beneficial to also include control conditions, which would allow for a more rigorous examination and to better understand which of the conditions is driving the effect.

In addition, we noticed some minor issues in the setup of the experiment, which may have impacted the results. First, the figures of the spinners described the outcome as ‘win’ and ‘lose’, whereas the target article described it as ‘win’ and ‘win nothing’. Although the two expressions may seem like they mean the same thing, they might carry different valence values and affect people’s decisions given classic gain-loss framing effects that impact risk-related decisions (Tversky & Kahneman, 1981). Second, Bown (2006) acknowledged that the pictorial presentation made tradeoffs of odds and payoffs between spinners more difficult, so it would have been better accompanied by a textual summary. As the lure of choice is a heuristic that stems from choice pairing, instead of a probability judgement error that arises from the failure to use the Bayes rule, it should be clear to participants what the relative differences are between the options.

## Bown et al. (2003): Design, hypotheses, and findings

We aimed to replicate all studies of the published target article[[1]](#footnote-2). We summarised the target article’s hypotheses in Table 1 and their findings in Table 2.

###### 

###### Table 1 *Target article’s Studies 1, 2, and 3: Summary of hypotheses*

| **Hypothesis** | **Studies involved** | **Scenarios** | **Description** |
| --- | --- | --- | --- |
| 1 | 1 | Nightclub  Bank | People are more likely to choose the target when it is paired with a lure. |
|  | 2 | Casino |  |
| 2 | 3 | Monty Hall | People are more likely to change their decisions when offered an opportunity to choose. |

*Note*. See Table 15 for mapping hypotheses, studies, scenarios, statistical tests, and effects comparisons between the target article and the replication.

###### Table 2 *Summary of effect size calculation and power analysis*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| DV | Our calculation | | | Original reports | | | Sample size required  (power = 0.95) |
|  | ***X****2* | ***Fisher’s p*** | **Cramer's *V* and *CI*** | ***X****2* | ***Fisher's p*** | **Cramer's *V*** |  |
| **Nightclub** |  |  |  |  |  |  |  |
| * Counts choosing Club Cherish and Diesel | 3.7 | .040 | .18 [0.01, 0.34] | 4.42 | < .05 | .18 | 391 |
| **Bank** |  |  |  |  |  |  |  |
| * Counts choosing Account A and B | 4.0 | .040 | .24 [0.03, 0.43] | 4.93 | < .05 | .24 | 225 |
| **Casino Dominated Lure** |  |  |  |  |  |  |  |
| * Counts choosing Spinner A and B | 1.1 | .300 | .11 [0.00, 0.28] | 1.56 | .46 | n.s. | / |
| **Casino Conflicted Lure** |  |  |  |  |  |  |  |
| * Counts choosing Spinner A and B | 6.4 | .007 | .27 [0.08, 0.43] | 7.45 | < .01 | .27 | 185 |
| **Monty Hall** |  |  |  |  |  |  |  |
| * Counts for sticking  and switching | 13 | < .001 | .19 [0.09, 0.28] | / | / | / | 438 |
| * 3-door contrast CAD | / | / | / | 2.62 | n.s. | .10 | / |
| * 3-door contrast CAC | / | / | / | 12.93 | < .001 | .24 | / |
| * CAD contrast CAC | / | / | / | 5.13 | < .05 | .14 | / |

Note: We used the raw data reported in Bown et al. (2003) for our calculation. Original reports are provided by Bown (2006). All calculations follow the third-way analysis (excluding choices of lures) reported by Bown et al. (2003). Please refer to the “Sample size before and after exclusions” subsection of the supplementary materials for comprehensive results, including and excluding choices of lures. X-square adjusted with Yates’ Continuity Correction (Hilton, 2005). Sample size calculations are based on our calculation.

## Extensions: Adding neutral no-grouping conditions

To extend the replication, we added neutral no-grouping conditions in which participants directly choose between the options without any pairing or two-stage choice. While the target article conducted pre-tests to ensure a roughly equal preference for the two targets, adding the no-grouping conditions provides a baseline for assessing the influence of partitioning on participants’ choices, and reveals their decision processes.

Extensive research has demonstrated that decision-making outcomes differ when alternatives are evaluated collectively compared to when they are evaluated individually (Kahneman & Ritov, 1994; Nowlis & Simonson, 1997; Hsee, 1996). Read et al. (1999) conducted a comprehensive review on the phenomenon of choice bracketing, which refers to the differences in choices made under narrow bracketing (individual consideration of options) versus broad bracketing (collective evaluation of options). The authors concluded that broad bracketing facilitates the emergence of adding-up effects, trade-offs, and other effects, which can fundamentally alter individuals' evaluations of alternatives. These concepts can also be applied in the context of the lure of choice. When individuals consider a larger choice set, they engage in broad bracketing by collectively evaluating one of the target options and the lure.

This broader evaluation leads to the overall benefits of the choice set, including the benefits of the target, the benefits of the lure, and the flexibility of choice, outweighing the benefits associated with selecting the isolated target option. Furthermore, when individuals are confronted with multiple choices involving multidimensional alternatives, they often strive to identify "integrative solutions." These integrative solutions involve combining the benefits of different options while compensating for their disadvantages. To illustrate, in the bank scenario, the lure option offers a shorter notice period compared to Account A, but a higher interest rate compared to Account B. In this case, the lure option serves as an integrative solution when paired with one of the target options. By compromising the benefits and costs of interest rates and notice periods, the choice set becomes more appealing than Account A or Account B individually. Therefore, the way options are grouped can lead to different tendencies in selecting choices from the same set of options.

If preferences in the experimental conditions do not differ from the no-grouping conditions, it suggests that people choose between all three options, that is they choose the outcome they want first and then follow the path that leads to the outcome. On the contrary, if preferences in the experimental conditions are different from the no-grouping conditions, it suggests that the grouping of options affects people’s decisions. In particular, it provides compelling evidence when more participants choose the choice set than the isolated item. Furthermore, as all the target options are superior to the lure, if more participants chose the target option than the lure, and the preference is different from the no-grouping conditions, it reveals that participants are lured towards the choice set without careful consideration of the characteristics of the options, but then compare the relative tradeoffs of the two viable options and make a decision.

Overall, by including the no-grouping conditions, our replication aimed to offer a comprehensive examination of the lure of choice phenomenon and provide insights into the underlying decision-making mechanisms of the lure of choice.

## Pre-registration and open-science

We provided all materials, data, and code on: <https://osf.io/e47jh/> . [To be updated in Stage 2:] 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, and exclusions conducted for this investigation are reported, and data collection was completed before analyses. This Registered Report was written using the Registered Report template by Feldman (2023).

# Method

[IMPORTANT: Method and results were written using a randomised 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 conducted a power analysis based on statistics reported in the target article. Effect size and power were all calculated with the help of a guide by Jané et al. (2024) and R (Version 4.3.2; R Core team, 2023) using packages “ufs” (Peters & Gruijters, 2023) and “pwrss" (Bulus, 2023). We focused on the differences between conditions that the authors identified as significant results, and aimed for a power of 0.95 with an alpha of 0.05. Additionally, we conducted a sensitivity test with GPower (Version 3.1; Faul et al., 2007).

We summarised the calculated required sample size for each study in Table 2, and the minimum required sample size was 438 participants in total.‎ This calculation is based on the effect size of the weakest effect of all the scenarios - Cramer’s V = 0.19 (Monty Hall scenario), with a power of 0.95, alpha of .05. We provided more information regarding these calculations in the “Power analysis of the target article effect to assess the required sample for replication” subsection of the supplementary materials.

Given the likelihood that the target article’s effects are overestimated, we used the “small-telescope” approach (Simonsohn, 2015) aiming for enough power to detect effects much weaker than those reported by the original study with the general rule of thumb (though meant for other similar designs) to multiply the initial required sample by 2.5. This resulted in a sample of 933. Accounting for possible exclusions of 10%, our integrated design, and allowing for the potential of additional analyses, we aimed for a larger total sample of 1026 participants, larger than the combined samples in the target article. A sensitivity analysis indicated that a sample of 155 per condition for a Fisher’s exact test allows detecting proportion differences of 0.30(95% power, alpha = 5%, two-tail), and that a sample of 933 (after exclusions) allows the detection for *w* = 0.16 for a six-conditions chi-squared goodness-of-fit test that takes into account our extension (95% power, alpha = 5%, two-tail). These are commonly considered small to moderate effects in social psychology (Jané et al., 2024).

[We tried our best to reproduce the original results, the code can be found in “Bown-etal-2003-analyses.Rmd/html”, yes Table 2 indicates some inconsistencies between our analyses and theirs. If possible, we would appreciate the reviewers’ help in improving on our analyses to better figure out how they achieved their reported results.]

## Participants

[Stage 1 note: To demonstrate what the results would look like after data collection we simulated a dataset of 1000 participants using Qualtrics and reported our analyses below based on that dataset. Results will later be updated in full to a sample of 1026 and the real data.]

We recruited a total of 1000 British participants on Prolific (*Mage* = 48.1, *SD* = 28.8; 255 females, 249 males, 496 others or did not disclose). We summarised a comparison of the target article samples and the replication samples in Table 3.

[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 analysed 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 £11.44/hour (April, 2024), per minute, so for example 5-8 minutes survey would be paid £1.6 per participant.]

[An example, to be updated in Stage 2: We first pretested survey duration with 30 participants to test time run estimate and adjusted pay based on the duration. The data of the 30 participants was not analysed other than to assess technical issues, survey completion duration, and needed pay adjustments, and were included in the final data analysis.]

###### Table 3 *Differences and similarities between target article and replication*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Bown et al. (2003) | | | | Replication |
|  | Study 1 Nightclub | Study 1  Bank | Study 2  Casino | Study 3  Monty Hall | UK Prolific |
| Sample size | 150  100 | 100 | 257 | 373 | 1000 | |
| Geographic origin | UK general public | UK general public | UK undergraduate students | UK undergraduate students | UK | |
| Gender | 48 males,  102 females | 36 males,  64 females | ~115 males,  ~142 females | ~205 males,  ~168 females | 249 males,  255 females,  496 other/did not disclose | |
| Median age (years) | Unreported | Unreported | Unreported | Unreported | 47.0 | |
| Average age (years) | 23.2 | 26.6 | 18.8 | 19.4 | 48.1 | |
| Standard deviation age (years) | Unreported | Unreported | Unreported | Unreported | 28.8 | |
| Age range (years) | Unreported | Unreported | Unreported | Unreported | 0-100 | |
| Medium (location) | Shopping centre | Shopping centre | University | University | Computer (online) | |
| Compensation | Unreported | Unreported | Unreported | Unreported | Nominal payment | |
| Year | 2003 or earlier | 2003 or earlier | 2003 or earlier | 2003 or earlier | 2024 | |

*Note*: Bown (2006) reported that 66% participants in Study 1 Nightclub, 64% participants in Study 1 Bank, 55% participants in Study 2 and 45% participants in Study 3 were female.

### Design: Replication and Extension

In the target article, Studies 1 (nightclub scenario and bank scenario combined), 2 (casino scenario) and 3 (Monty Hall problem) were conducted separately with independent samples. We ran the four scenarios together in a single unified data collection. The display order of scenarios was counterbalanced using the randomizer “evenly present” function in Qualtrics. 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), and 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.]

We summarised the experimental design in Tables 4 to 7, and our adjustments to the target article in Table 9. In Studies 1 and 2, the floating lure design was employed, where participants were presented with a choice between two targets and a lure. Both targets served as targetI and targetL for different conditions.

The design of the nightclub scenario was a 2 (targetL: Club Diesel and Club Cherish; between-subjects) by 2 (lure direction: north and south; between-subjects design). Similarly, the design of the Bank scenario was a 2 (targetL: Account A and Account B; between-subjects) by 2 (bank: bank J and bank K; between-subjects design). The design of the Casino scenario was a 2 (targetL: Spinner A and Spinner B; between-subjects) by 2 (lure: dominated lure and conflicted lure; between-subjects design). Lastly, for Study 3, the Monty Hall problem followed a between-subject design with a single factor of condition (3-door, 4-door choose-a-door, and 4-door choose-a-choice; between-subjects design).

###### Table 4 *Study 1 Nightclub scenario: Replication and extension experimental design (between-subject)*

|  |  |  |  |
| --- | --- | --- | --- |
| IV1:TargetL (main)  IV2:Lure direction (secondary) | **IV1: TargetL = Club Diesel**  Diesel paired with lure (Atom)  Direction X:  Cherish  Direction Y:  Diesel  Atom | **IV1: TargetL = Club Cherish**  Cherish paired with lure (Atom)  Direction X:  Diesel  Direction Y:  Cherish  Atom | **IV1: No TargetL [Extension]**  No grouping  (randomised choice order) |
| **IV2: North**  Lure and TargetL in north  **IV2: South**  Lure and TargetL in south | DV1: Choice of direction  (direction with one nightclub/ direction with two nightclubs)  DV2: Choice of nightclub (Diesel / Cherish / Atom [lure]) | | |
|  |  | |  |

###### Table 5 *Study 1 Bank scenario: Replication and extension experimental design (between-subject)*

|  |  |  |  |
| --- | --- | --- | --- |
| IV1:TargetL (main)  IV2:Lure bank (secondary) | **IV1: TargetL = Account A**  Account A paired with lure (Account L)  Bank X:  Account B  Bank Y:  Account A  Account L | **IV1: TargetL = Account B**  Account B paired with lure (Account L)  Bank X:  Account A  Bank Y:  Account B  Account L | **IV1: No TargetL [Extension]**  No grouping  (randomised choice order) |
| **IV2: Bank J**  Lure and TargetL in Bank J  **IV2: Bank K**  Lure and TargetL in Bank K | DV1: Choice of bank  (one-account bank/ two-account bank)  DV2: Choice of account  (Account A/ B/ L [lure]) | |  |
|  |  | |  |

###### 

###### Table 6 *Study 2 Casino scenario: Replication and extension experimental design (between-subject)*

|  |  |  |  |
| --- | --- | --- | --- |
| IV1: TargetL (main)  IV2: Lure (secondary) | **IV1: TargetL = Spinner A**  Spinner A paired with lure  Table 1:  Spinner A  Spinner D/C  Table 2:  Spinner B | **IV1: TargetL = Spinner B**  Spinner B paired with lure  Table 1:  Spinner B  Spinner D/C  Table 2:  Spinner A | **IV1: No TargetL [Extension]**  No grouping  (randomised choice order) |
| **IV2: Dominated lure**  Lure (Spinner D) is weakly dominated by targets | DV1: Choice of table  (one-spinner table/ two-spinner table)  DV2: Choice of spinner  (Spinner A/ B/ C [lure] / D [lure]) | |  |
| **IV2: Conflicted lure**  Lure (Spinner C) is conflicted with targets |

###### 

###### Table 7 *Study 3 Monty Hall scenario: Replication and extension experimental design (between-subject)*

|  |  |  |
| --- | --- | --- |
| **IV1: 3 Door**  Stick or switch to another door | **IV1: 4 door Choose-A-Door**  Stick or switch to one of two other doors | **IV1: 4 door Choose-A-Choice**  Stick or switch to one of two other doors, which to be decided later |
| DV1: Choice of stick or switch  DV2: Choice of door [Extension]  (Door K/ L/ M) | | |
|  | | |

###### Table 8 *Position of choice and lure across scenarios*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Study** | **Choice** | **Lure** |
| **Nightclub** | 1 | Up / down | Left |
| **Bank** | 1 | Left / right | Right |
| **Casino** | 2 | Left / right | Bottom Left |
| **Monty Hall** | 3 | Left / right | N/A |

## 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/41a2e5fa-1056-4ca2-b86c-75b5dfd00735/SV_3EsmubjrRPQRpVI?Q_CHL=preview&Q_SurveyVersionID=current> ]

We reconstructed the target’s stimuli and adjusted it to an online Qualtrics survey based on the information provided in the article and the thesis it is based on.

Participants first indicated 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 order of the options being rotated (yes, no, not sure). They were then randomly assigned into the experimental conditions, which were counterbalanced using the randomizer ‘evenly present’ function in Qualtrics, and responded to each assigned condition for a total of four scenarios.

### Nightclub scenario

We summarised the design and choices in Table 4. The differences between the nightclubs are summarised in Figure 1.

In the nightclub scenario, participants were asked to select a nightclub out of three, each varied in the entrance fee and whether they played the participants’ favourite music. Club Atom (lure) was dominated by Club Diesel. The scenario included six conditions: “Cherish North” (*n* = 166), “Cherish South” (*n* = 166), “Diesel North” (*n* = 167), “Diesel South” (*n* = 167), “No-grouping CD” (*n* = 167), and “No-grouping DC” (*n* = 167). The first four conditions were replications of the target article, whereas the last two were our neutral control conditions.

In the Cherish conditions, Club Atom (lure) was paired with Club Cherish, while Club Diesel was presented in isolation. Conversely, in the Diesel conditions, Club Atom was paired with Club Diesel, while Club Cherish was presented in isolation. The location of the targets varied based on the condition. In the North conditions, both targetL and the lure were located in the north, while in the South conditions, both targetL and lure were located in the south. Participants first decided on the direction. Those choosing the direction with two clubs then chose the nightclub they preferred, whereas those choosing the direction with only one club confirmed their choice.

We added two no-grouping conditions, where participants directly chose their preferred nightclub without any pairings or directions. We randomised the display order of the three nightclubs across the two conditions. In the “No-grouping CD” condition, Club Cherish was presented before Club Diesel, whereas in the “No-grouping DC” condition, Club Diesel was presented before Club Cherish. In both no-grouping conditions, Club Atom (lure) was consistently presented last.

### Bank scenario

We summarised the design and choices in Table 5. The differences between the accounts are summarised in Figure 2.

In the bank scenario, participants were asked to imagine that they had inherited some money and had to choose a savings account. They were presented with three savings accounts in two banks, with one bank offering two possible accounts, and the other offering one. Account L (lure) conflicted with both targets (better on one dimension and worse on the other).

The scenario included six conditions: “Account A with lure in Bank J” (*n* = 167), “Account B with lure in Bank J” (*n* = 166), “Account A with lure in Bank K” (*n* = 167), “Account B with lure in Bank K” (*n* = 167), “No-grouping AB” (*n* = 166), and “No-grouping BA” (*n* = 167). The first four conditions were replications of the target article, while the last two were our neutral control conditions. In Account A conditions, Account L (lure) was paired with Account A, while Account B was presented in isolation. Conversely, in Account B conditions, Account L was paired with Account B, while Account A was presented in isolation. The bank offering the targets varied based on the condition. In Bank J conditions, targetL and the lure were provided by Bank J, while in Bank K conditions, targetL and lure were provided by Bank K. Participants first decided which bank they were going to visit. Those selecting the two-account bank then chose the account they preferred, while those selecting the one-account bank then confirmed their choice of account.

We added two no-grouping conditions where participants directly choose their preferred accounts without any pairings or banks. We randomised the order of the account presentation across the two conditions. In “No-grouping AB”, Account A was presented before Account B, whereas in “No-grouping BA”, Account B was presented before Account A. In both no-grouping conditions, Account L (lure) was consistently presented last.

### Casino scenario

We summarised the design and choices in Table 6. The differences between the spinners are summarised in Figure 3.

In the casino scenario, participants were asked to imagine themselves in a casino with a token to spend on a roulette-type game. They were presented with three options: Spinner A, Spinner B (both targets), and a lure. The expected winnings of both targets were £45, while the dominated lure (Spinner D) had an expected winning of £37.5 and the conflicted lure (Spinner C) had an expected winning of £37.6.

The study included five conditions: “A and Dominated” (*n* = 199), “A and Conflicted” (*n* = 201), “B and Dominated” (*n* = 200), “B and Conflicted” (*n* = 200), and “No-grouping” (*n* = 200). All conditions were replications of the target article, yet we modified the “No-grouping” condition.

The spinners denoted in the condition names were placed on the left table, while the remaining target was placed on the right table. In the “A and Dominated” condition, Spinner A and D were on the left table, leaving Spinner B presented in isolation on the right table. In the “A and Conflicted” condition, Spinner A and C were on the left table, leaving Spinner B presented in isolation on the right table. In the “B and Dominated” condition, Spinner B and D were on the left table, leaving Spinner A presented in isolation on the right table. In the “B and Conflicted” condition, Spinner B and C were on the left table, leaving Spinner A presented in isolation on the right table.

Participants first chose the table they would like to spend the token. Those selecting the two-spinner table then choosing their preferred spinner, whereas those selecting the one-spinner table confirmed their choice of spinner. Instead of having four separate control conditions with different groupings of spinners, we included all four spinners in a single control condition, isolating the effects of pairings.

### Monty Hall problem scenario

We summarised the design and choices in Table 7. Examples of stimuli were provided in Figures 4 and 5.

The Monty Hall problem consisted of three conditions: “3-door” (*n* = 333), “4-door choose-a-door” (*n* = 334), and “4-door choose-a-choice” (*n* = 333). In the traditional “3-door” condition, participants were shown three doors, one of which hid a star prize Ferrari, while the others had goats as consolation prizes. Participants were told that they initially chose Door K, and one of the other doors without the Ferrari was opened. Participants could then choose to stick with Door K or switch to Door L. In 4-door conditions, participants were shown four doors. In the “4-door choose-a-door” condition, participants could stick with Door K, switch to Door L or switch to Door M. In the “4-door choose-a-choice” condition, participants could either stick with Door K or choose to switch to Door L or Door M, which they did not have to make their final decision immediately but would indicate their choice on the subsequent page. As an extension, we also tested participant’s familiarity with the Monty Hall problem and their opinion on the optimal solution.

At the end of the experiment, participants answered a number of funnelling and demographic questions, and were debriefed. We provided a more comprehensive overview of the survey procedure and complete instructions to participants in “Materials and scales used in the replication + extension experiment” in the supplementary.

## Deviations

We made a few adjustments to the design of the target article and summarised those in Table 9.

###### Table 9 *Replication and extension adjustments to the target article’s methods and design*

| **Studies** | **Factor** | **Target article** | **Adjustment in current study** | **Reason for change / Justifications** |
| --- | --- | --- | --- | --- |
| All | Study design | Participants completed the studies with pen and paper in a shopping centre (study 1), university (study 2), and an interactive game setting in class (study 3) | Participants completed the studies on an online survey | Lower cost, higher efficiency, more diverse population (generalizability) |
| All | Procedures | The four scenarios are conducted separately | All scenarios are conducted in one data collection | Address issues with sample/context, allow for insights comparing scenarios with the same participants |
| All | Sample size | *N =* 880 (3 studies combined) | *N =* 1000 | Based on a power analysis |
| All | Stimuli set | No headings and instructions | Include headings and instructions (e.g., imagine you enjoy going to nightclubs) | Better flow, clearer instructions, provides more detailed context |
| All | Stimuli set | No reminder of the scenario | Include a reminder of the scenario in subsequent pages of the study | Reduce cognitive load and remind participants information about the options when they make the choice |
| 1 | Stimuli set | The descriptions of options in the target article were different from the stimuli actually used  E.g., Club Cherish is in the north of the town - the entrance fee is £4, they do not really play your favourite kind of music but you are likely to bump into other friends there.  Note that Bown et al. (2003) wrote ‘This club is cheap (£4) but does not play very good music’  Note that Bown et al. (2003) wrote ‘instant access for withdrawals’ but the stimuli wrote ‘instant access to your money should you need it’ | Aligned all descriptions to the original stimuli, and deleted extra words that may distract participants.  E.g., Club Cherish is in the north of the town - the entrance fee is £8, they do not play your favourite kind of music. | Price adjusted according to inflation rate (doubled).  Removed ‘really’ to keep descriptions neutral.  Removed ‘but you are likely to bump into other friends there’ because the statement is not a differentiating factor, and we aim to keep participants focused on important factors.  Changed categorization of ‘cheap’ ‘expensive’ ‘moderately expensive’ to prices given that labels may cause bias. |
| 1, 2 | Stimuli set | Provided only the options (e.g., go North/ go South) | Options also specify the number of choices in the options (e.g., Go North with two nightclubs (which you will choose in the next screen)/Go South with one nightclub) | Give enough information to participants to ensure there is no confusion and they know what they are choosing without having to reread the scenarios |
| 1, 2 | Stimuli set | Short descriptions, such as - “If you chose to go North, which nightclub would you go to?” | Longer clearer descriptions, such as “You asked the taxi driver to go North and now face a decision between the 2 clubs, which nightclub would you go to? (for the two-choice option)”  OR  “You asked the taxi driver to go South, which only has 1 club, confirm your choice of the 1 club in that area by checking the box. (for the one-choice option)” | Include the number of options of the choice made and confirmation for the one-choice option to ensure participants understood the scenarios and choices they made, with a confirmation of their choice, even for the direction with the 1 option. |
| 1, 2 | Stimuli set | Options indicated only the name of the options in final choice (e.g., Club Atom/ Club Cherish) | Options include information about option features in the final choice (e.g., North: **Club Cherish**: the entrance fee is £8, they do not play your favourite kind of music.) | Give as much information to participants to ensure there is no confusion and they know what they are choosing without having to reread the scenario |
| 2 | Stimuli set | The position of each table (left or right of the page) and the spinners on each table (top or bottom) were fully counterbalanced | Present two spinners (TargetL and lure) on Table 1 and another spinner (TargetI) on Table 2 | Target article found no indication for position effects. Given we ran many scenarios, we ensured positions of choices and lures were varied on the scenario level (see Table 8) |
| 2 | Stimuli set | Only present odds in images of the spinners | Included also text odds in the images, descriptions of the scenario, and options. | Give as much information to participants to ensure there is no confusion and they know what they are choosing without having to reread the scenario |
| 2 | Stimuli set | The stimuli wrote ‘lose’ in the spinner but Bown et al. (2003) wrote ‘winning nothing’ | Presented ‘win’ and ‘win nothing’ | Align description and avoid framing effects |
| 2 | Stimuli set | In Bown et al. (2003), lure C did not detail the chances of “winning nothing”  (Lure C: 47% chance of £40) | Include the odds of both “winning” and “winning nothing” for all spinners  e.g., Lure C: 47% chance of winning £70 (and 53% chance of winning nothing) | Provide comprehensive information to ensure participants make informed decisions |
| 3 | Procedure | The original study asked participants to explain explicitly the reasons behind their decision.  E.g., “In a sentence of two, please tell us why you chose the way you did. That is, to either stick with box K or to change to one of the other two boxes” | We asked the participants whether they had heard of the Monty Hall problem, what they thought the optimal solution was, and the reason behind it | We anticipated that familiarity with the Monty Hall problem would be one of the major reasons causing participants to switch, and the guided questions are easier to analyse and more friendly to the participants |
| 3 | Statistical analyses | Comparing the conditions pairwise with three Fisher’s exact tests | Compared the three conditions with chi-square post hoc tests | Enabling a comprehensive analysis across the three conditions |
|  |  |  |  |  |

## Evaluation criteria for replication findings

We aimed to compare the replication effects with the original effects in the target article using the criteria set by LeBel et al. (2019).

We pre-registered our overall strategy to conclude a successful replication if 3 or 4 out of the 4 scenarios showed a signal in the same direction as the target article, a failed replication if no scenario showed a signal in the same direction as the target, and 1 or 2 out of the 4 scenarios to be a mixed results replication. Study 2 includes two tests and will be considered successful if both are supported, mixed if only 1, and failed if both are not supported. We summarised all hypotheses per scenarios in Table 15.

## Replication closeness evaluation

We provided details on the classification of the replications using the criteria by LeBel et al. (2018) criteria in Table 10 below (see section “replication closeness evaluation” in the supplementary). We summarised the replication as a "close” replication.

###### Table 10 *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 | Similar | The original article systematically varied the positions of the targets but found no indication for position effects. We decided to hold it constant in Study 2 and vary across scenarios.  We included extension control conditions in the data collection (which should not impact the replication conditions). |
| DV construct | Same |  |
| IV operationalization | Same |  |
| DV operationalization | Same |  |
| IV stimuli | Similar | We adjusted the options and scenarios (see deviations). |
| DV stimuli | Similar | We added confirmation for participants who chose the one-option choice in Studies 1 and 2. |
| Procedural details | Similar | We ran all scenarios in a unified single data collection with randomised scenario presentation order |
| Physical settings | Similar | Experiment was conducted online instead of via traditional paper survey |
| Contextual variables | Different | Participants were recruited online using Prolific. |
| Population (e.g., age) | Similar | Participants were of the same ethnicity but generally older based on the simulated data |
| Replication classification | Close 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, we first ran Fisher’s exact tests to examine the difference between conditions for each scenario. Fisher's exact tests are useful for analysing categorical data when the sample sizes are small. We also added chi-square tests to cater for our larger sample size. We adjusted our Fisher’s *p*-value and chi-square *p*-value with the false discovery rate correction (Benjamini & Hochberg, 1995) to minimise false discoveries in multiple comparisons.

In Studies 1 and 2, we followed the target article's analyses by excluding choices of lures from the data. This exclusion enabled a focused examination of how the presence or absence of the lure affects the number of participants choosing each target.

In Study 3, instead of conducting three pairwise Fisher's exact tests to examine the differences across the three conditions, we opted for a single Fisher's exact test and chi-square test. This approach facilitated a more comprehensive analysis of the differences among the conditions. Additionally, post-hoc tests were employed to further explore the specific contrasts between the conditions.

### Extensions

In addition to the replication conditions, we performed additional analyses that combined the control conditions and replication conditions in a single analysis. By incorporating the control conditions, we aimed to compare and contrast the effects observed in the replication conditions with those in a controlled setting.

In Study 3, we included a measure to inquire about participants' familiarity with the Monty Hall problem. This measure was added as an exploratory component and will not be subjected to formal analysis. Its purpose was to capture participants' prior knowledge or exposure to the Monty Hall problem in case our predictions were not supported by the quantitative data. By including this measure, we could gain additional insights into participants' background understanding of the Monty Hall problem in case hypotheses are not supported, potentially providing context for their responses and facilitating the interpretation of the study results.

### Outliers and exclusions

We will run our analyses on the full sample of all those who completed the study successfully and answered all questions. Those who dropped out will not be included. We do not plan on identifying outliers in the data analysis for this study. All data collected from participants who completed the study will be included in the analysis.

# Simulated Results

[IMPORTANT: Method and results were written using a randomised 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 summarised the descriptives in Tables 11 to 14 and statistical tests in Tables 15 and 16, with plots in Figures 6 to 13. We conducted a Fisher’s exact test and a chi-square test for each scenario with the four replication conditions after excluding choices of lures. The following analyses were all performed with R (Version: 4.3.2) with packages "pwr", "ufs", "jmv" (Selker et al., 2023), “rcompanion” (Mangiafico, 2023), “tidyverse” (Wickham et al., 2019), "ggstatsplot" (Patil, 2021), "rstatix" (Kassambara, 2023).

## Replication

In the nightclub scenario, we found support for differences in the number of participants choosing each target nightclub between the “Cherish North” condition (122 out of 166, 73%), “Cherish South” condition (120 out of 166, 72%), “Diesel North” condition (135 out of 167, 81%), “Diesel South” condition (120 out of 167, 72%) after excluding choices of lures X2(3) = 49.74, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .32, 95% CI [0.23, 0.39]. Specifically, more participants chose Club Cherish in the “Diesel North” condition (88 out of 167, 53%) than in the “Cherish North” condition (50 out of 166, 30%), whereas more participants chose Club Diesel in the “Cherish North” condition (72 out of 166, 43%) than in the “Diesel North” condition (47 out of 167, 28%; adj. *chisq* *p* < .001, adj. *fisher’s p* < .001). We provided a summary plot in Figure 6. In comparison, the target article reported X2(1) = 4.42, *fisher’s p* = < .05; *V* = .18, 95% CI [0.01, 0.34].

In the bank scenario, we found support for a difference in the number of participants choosing each target account between “Account A with lure in Bank J” condition (125 out of 167, 75%), “Account B with lure in Bank J” condition (123 out of 166, 74%), “Account A with lure in Bank K” condition (124 out of 167, 74%), and “Account B with lure in Bank K” condition (124 out of 167, 74%) after excluding choices of lures X2(3) = 57.44, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .34, 95% CI [0.26, 0.42]. Specifically, more participants chose Account A in the “Account B with lure in Bank J” condition (85 out of 166, 51%) than in the “Account A with lure in Bank J” condition (40 out of 167, 24%), whereas more participants chose Account B in the “Account A with lure in Bank J” condition (85 out of 167, 51%) than in the “Account B with lure in Bank J” condition (38 out of 166, 23%; adj. *chisq* *p* < .001, adj. *fisher’s p* < .001). Also, more participants chose Account A in the “Account B with lure in Bank K” condition (81 out of 167, 49%) than in the “Account A with lure in Bank K” condition (43 out of 167, 26%), whereas more participants chose Account B in the “Account A with lure in Bank J” condition (81 out of 167, 49%) than in the “Account B with lure in Bank K” condition (43 out of 167, 26%). We provided a summary plot in Figure 8. In comparison, the target article reported X2(1) = 4.93, *fisher’s p* < .05; *V* = .24, 95% CI [0.03, 0.43].

In the casino scenario, for dominated lure conditions, we found support for differences in the number of participants choosing each target spinner between the “A and Dominated” condition (142 out of 199, 71%) and the “B and Dominated” condition (143 out of 200, 72%) after excluding lures X2(1) = 34.39, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .35, 95% CI [0.24, 0.45]. Specifically, more participants chose Spinner A in the “B and Dominated” condition (98 out of 200, 49%) than in the “A and Dominated” condition (48 out of 199, 24%), whereas more participants chose Spinner B in the “A and Dominated” condition (94 out of 199, 47%) than in the “B and Dominated” condition (45 out of 200, 23%; adj. *chisq* *p* < .001, adj. *fisher’s p* < .001). We provided a summary plot in Figure 10. In comparison, the target article reported X2(1) = 1.56, *fisher’s p* = .46; *V* = .11, 95% CI [0.00, 0.28].

For conflicted lure conditions, we found support for differences in the number of participants choosing each target spinner between the “A and conflicted” condition (149 out of 201, 74%) and the “B and conflicted” condition (151 out of 200, 76%) after excluding lures X2(1) = 18.29, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .25, 95% CI [0.14, 0.35]. Specifically, more participants chose Spinner A in the “B and Conflicted” condition (92 out of 200, 46%) than in the “A and Conflicted” condition (54 out of 201, 27%), whereas more participants chose Spinner B in the “A and Conflicted” condition (95 out of 201, 47%) than in the “B and Conflicted” condition (59 out of 200, 30%). We provided a summary plot in Figure 11. In comparison, the target article reported X2(1) = 7.45, *fisher’s p* < .001; *V* = .27, 95% CI [0.08, 0.43].

In the Monty Hall problem scenarios, we found support for differences in the number of participants sticking to the original door or switching to another door between the “3-door” condition (*n* = 333), “4-door choose-a-door” condition (*n* = 334), and “4-door choose-a-choice” condition (*n* = 333), X2(2) = 19.40, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .14, 95% CI [0.08, 0.20]. Specifically, more participants stuck in the “3-door” condition (171 out of 333, 51%) than in the “4-door choose-a-door” condition (117 out of 334, 35%), whereas more participants switched in the “4-door choose-a-door” condition (217 out of 334, 65%) than in both the “3-door” condition (162 out of 333, 49%; adj. *chisq* *p* < .001, adj. *fisher’s p* < .001), and the “4-door choose-a-choice” condition (176 out of 333, 53%; adj. *chisq* *p* = .002, adj. *fisher’s p* = .003). We provided a summary plot in Figure 13. In comparison, the target article found X2(2) = 13, *fisher’s p* < .001; *V* = .19, 95% CI [0.09, 0.28].

###### Table 11 *Study 1 Nightclub scenario (with control conditions including lures): Descriptive statistics*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Condition/  Choice of club | Cherish North  (n = 166) | Cherish South  (n = 166) | Diesel North  (n = 167) | Diesel South  (n = 167) | No-grouping CD (n = 167) | No-grouping DC (n = 167) | Overall  (n = 1000 ) |
| Cherish | 50 (30.12%) | 34 (20.48%) | 88 (52.69%) | 78 (46.71%) | 52 (31.14%) | 54 (32.34%) | 356 (35.60%) |
| Diesel | 72 (43.37%) | 86 (51.81%) | 47 (28.14%) | 42 (25.15%) | 66 (39.52%) | 49 (29.34%) | 362 (36.20%) |
| Atom (lure) | 44 (26.51%) | 46 (27.71%) | 32 (19.16%) | 47 (28.14%) | 49 (29.34%) | 64 (38.32%) | 282 (28.20%) |

*Note*. The numbers denote counts of participants selecting each option (including lures), percentages are shown in parentheses. n indicates the sample size for that condition.

###### Table 12 *Study 1 Bank scenario (with control conditions including lures): Descriptive statistics*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Condition/  Choice of account | Account A with lure in Bank J  (n = 167) | Account B with lure in Bank J  (n = 166) | Account A with lure in Bank K  (n = 167) | Account B with lure in Bank K  (n = 167) | No-grouping AB (n = 166) | No-grouping BA (n = 167) | Overall  (n = 1000 ) |
| Account A | 40 (23.95%) | 85 (51.20%) | 43 (25.75%) | 81 (48.50%) | 56 (33.73%) | 62 (37.13%) | 367 (36.70%) |
| Account B | 85 (50.90%) | 38 (22.89%) | 81 (48.50%) | 43 (25.75%) | 52 (31.33%) | 47 (28.14%) | 346 (34.60%) |
| Account L (lure) | 42 (25.15%) | 43 (25.90%) | 43 (25.75%) | 43 (25.75%) | 58 (34.94%) | 58 (34.73%) | 287 (28.70%) |

*Note*. The numbers denote counts of participants selecting each option (including lures), percentages are shown in parentheses. N indicates the sample size for that condition.

###### Table 13 *Study 2 Casino scenario (with control conditions including lures): Descriptive statistics*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Condition/  Choice of account | A & Dominated  (n = 199) | A & Conflicted  (n = 201) | B & Dominated  (n = 200) | B & Conflicted  (n = 200) | No-grouping  (n = 200) | Overall  (n = 1000 ) |
| Spinner A | 48 (24.12%) | 54 (26.87%) | 98 (49.00%) | 92 (46.00%) | 49 (24.50%) | 341 (34.10%) |
| Spinner B | 94 (47.24%) | 95 (47.26%) | 45 (22.50%) | 59 (29.50%) | 48 (24.00%) | 341 (34.10%) |
| Spinner C (lure) | N/A | 52 (25.87%) | N/A | 49 (24.50%) | 50 (25.00%) | 151 (15.10%) |
| Spinner D (lure) | 57 (28.64%) | N/A | 57 (28.50%) | N/A | 53 (26.50%) | 167 (16.70%) |

*Note*. The numbers denote counts of participants selecting each option (including lures), percentages are shown in parentheses. N indicates the sample size for that condition.

###### Table 14 *Study 3 Monty Hall scenario: Descriptive statistics*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Condition/  Choice of stick or switch | 3-door  (n = 333) | 4-door CAD  (n = 334) | 4-door CAC  (n = 333) | Overall  (n = 1000 ) |
| Stick | 171 (51.35%) | 117 (35.03%) | 157 (47.15%) | 445 (44.50%) |
| Switch | 162 (48.65%) | 217 (64.97%) | 176 (52.85%) | 555 (55.50%) |

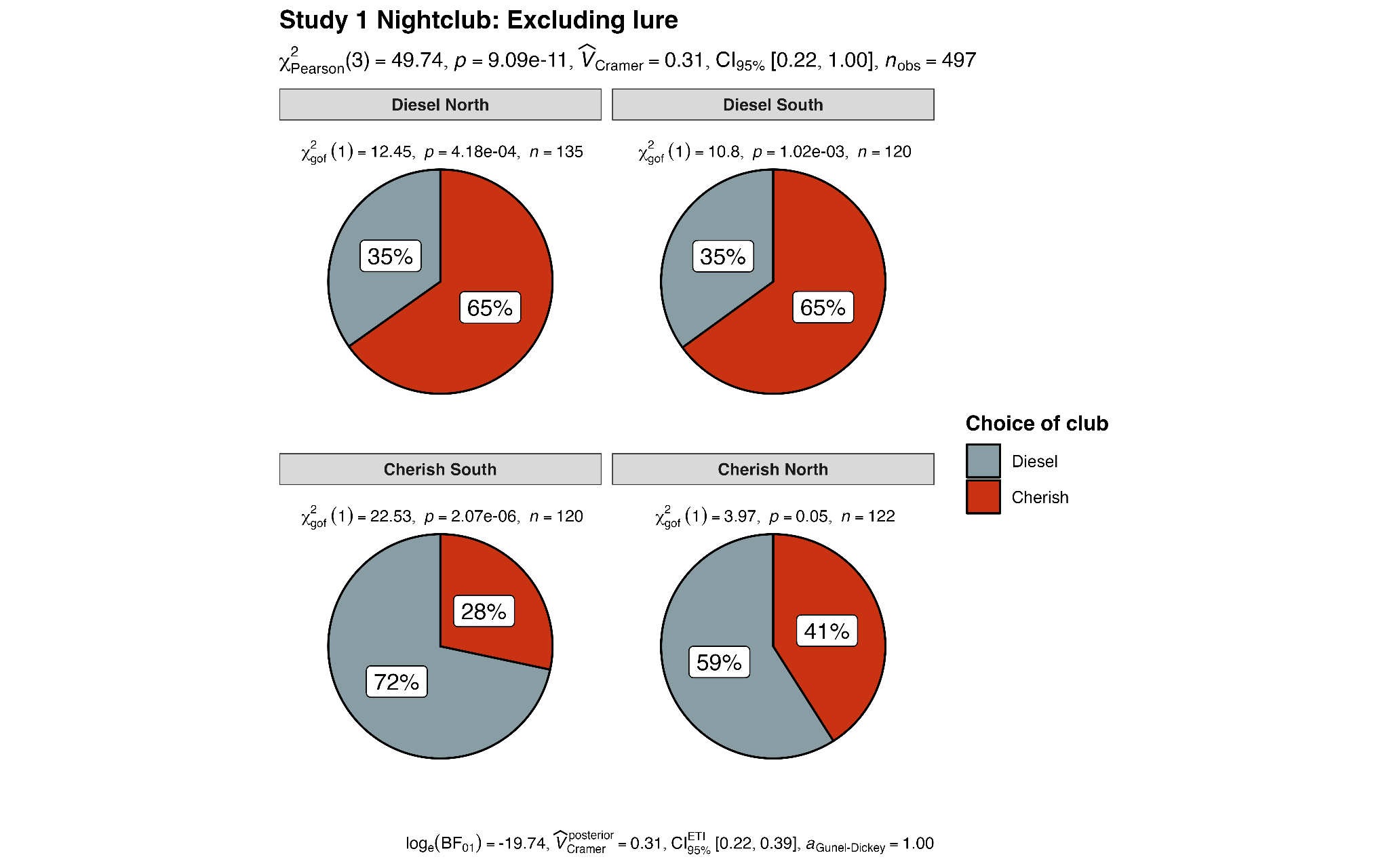
*Note*. The numbers denote counts of participants selecting each option (including lures), percentages are shown in parentheses. N indicates the sample size for that condition.

###### Table 15 *Replication: Summary of statistical tests, effects, and evaluation of replication findings against the target article*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Scenario | Replication | | | | | Target article | | | |  |
|  | ***df*** | ***X****2* | ***Chisq’s p*** | ***Fisher’s p*** | **Cramer's *V* and *CI*** | ***df*** | ***X****2* | ***Fisher's p*** | **Cramer's *V* and *CI*** | **Interpretation** |
| **Study 1: Nightclub [H1]** | **3** | **49.74** | **< .001** | **< .001** | **.32 [0.23, 0.39]** | **1** | **4.42** | **< .05** | **.18 [0.01, 0.34]** | **Signal, consistent, larger** |
| * With control conditions | 5 | 51.61 | < .001 | < .001 | .27 [0.20, 0.33] | / | / | / | / | Signal |
| **Study 1: Bank [H1]** | **3** | **57.44** | **< .001** | **< .001** | **.34 [0.26, 0.42]** | **1** | **4.93** | **< .05** | **.24 [0.03, 0.43]** | **Signal, consistent, larger** |
| * With control conditions | 5 | 59.09 | < .001 | < .001 | .29 [0.22, 0.35] | / | / | / | / | Signal |
| **Study 2:  Casino Dominated lure [H1]** | **1** | **34.39** | **< .001** | **< .001** | **.35 [0.24, 0.45]** | **1** | **1.56** | **.46** | **.11 [0, 0.28]** | **Signal, inconsistent, larger** |
| **Study 2: Casino Conflicted lure [H1]** | **1** | **18.29** | **< .001** | **< .001** | **.25 [0.14, 0.35]** | **1** | **7.45** | **< .01** | **.27 [0.08, 0.43]†** | **Signal, consistent, smaller** |
| Study 2:Casino [H1]   * With control condition | 4 | 53.05 | < .001 | < .001 | .28 [0.21, 0.35] | / | / | / | / | Signal |
| **Study 3: Monty Hall [H2]** | **2** | **19.40** | **< .001** | **< .001** | **.14 [0.08, 0.20]** | **2** | **13** | **< .001** | **.19 [0.09, 0.84]†** | **Signal, consistent, smaller** |

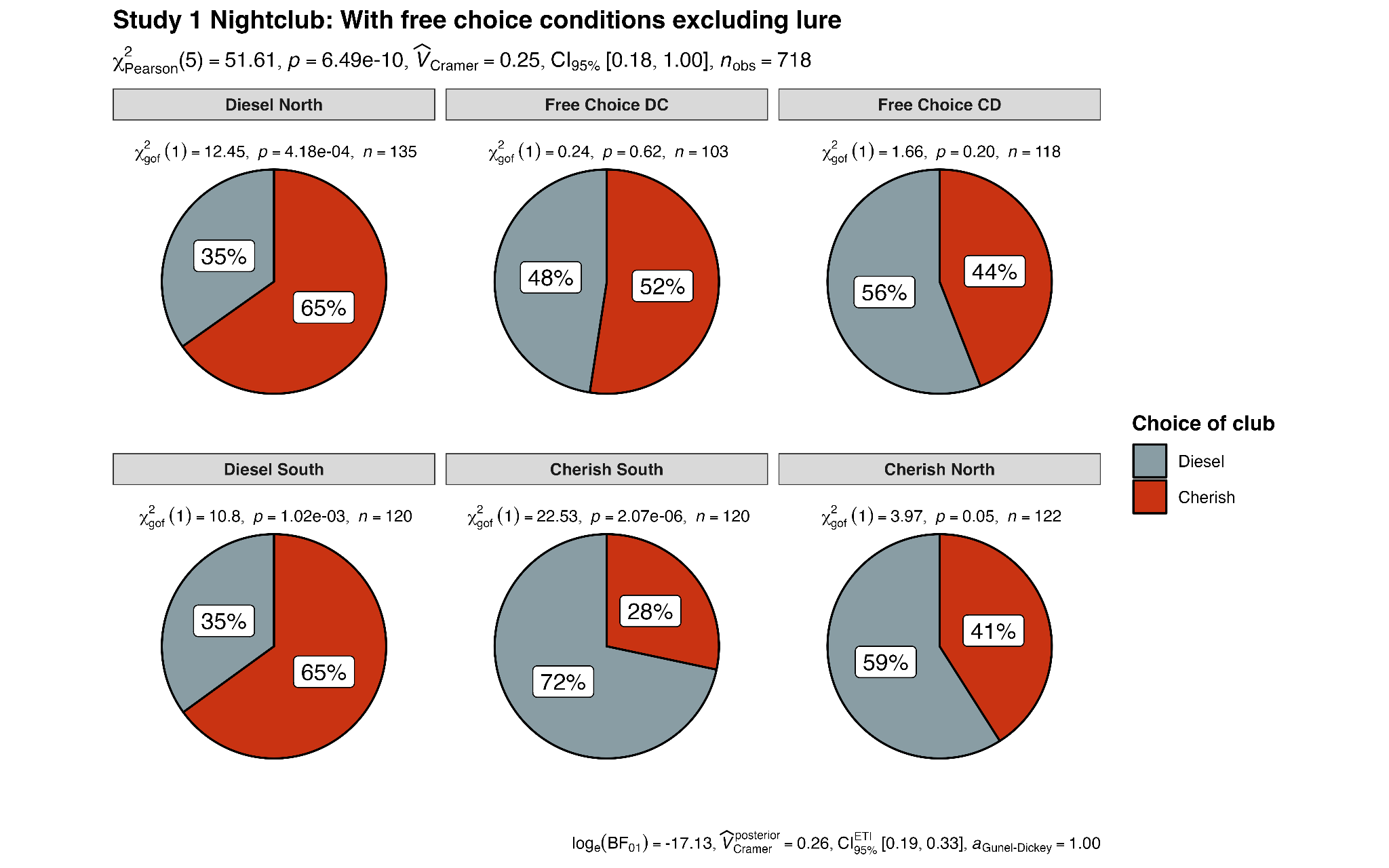
*Note*. All results are based on data after excluding lures. Fisher’s exact test. *CI* = 95% confidence intervals. The interpretation of the outcome was based on LeBel et al. (2019). Bolded indicates replication. No chi-square tests were done in the target article. †Cramer’s V was based on our calculation and was different from what the original article reported.   
Evaluation criteria (number of scenarios our of the 4 scenarios, with signal in the same direction as the target):  
Successful: 3 or 4; Failed: no scenario; Mixed: 1 or 2.  
Study 2 includes two tests and will be considered successful if both are supported, mixed if only 1, and failed if both are not supported.

###### Figure 6 *Study 1 Nightclub scenario: Comparison of conditions (excluding lure)*



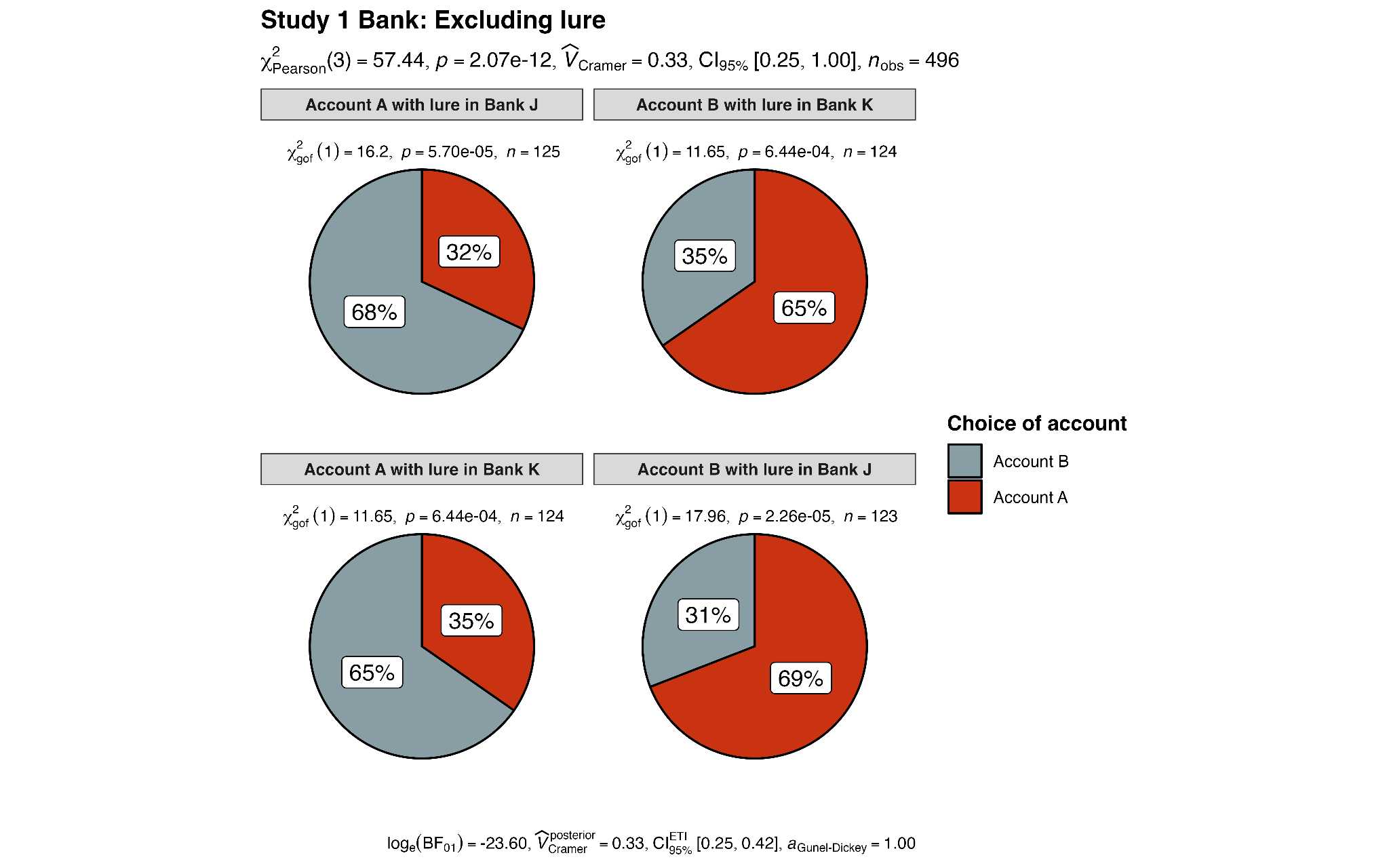
*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 7 *Study 1 Nightclub scenario: Comparison between replication and no-grouping conditions (excluding lure)*



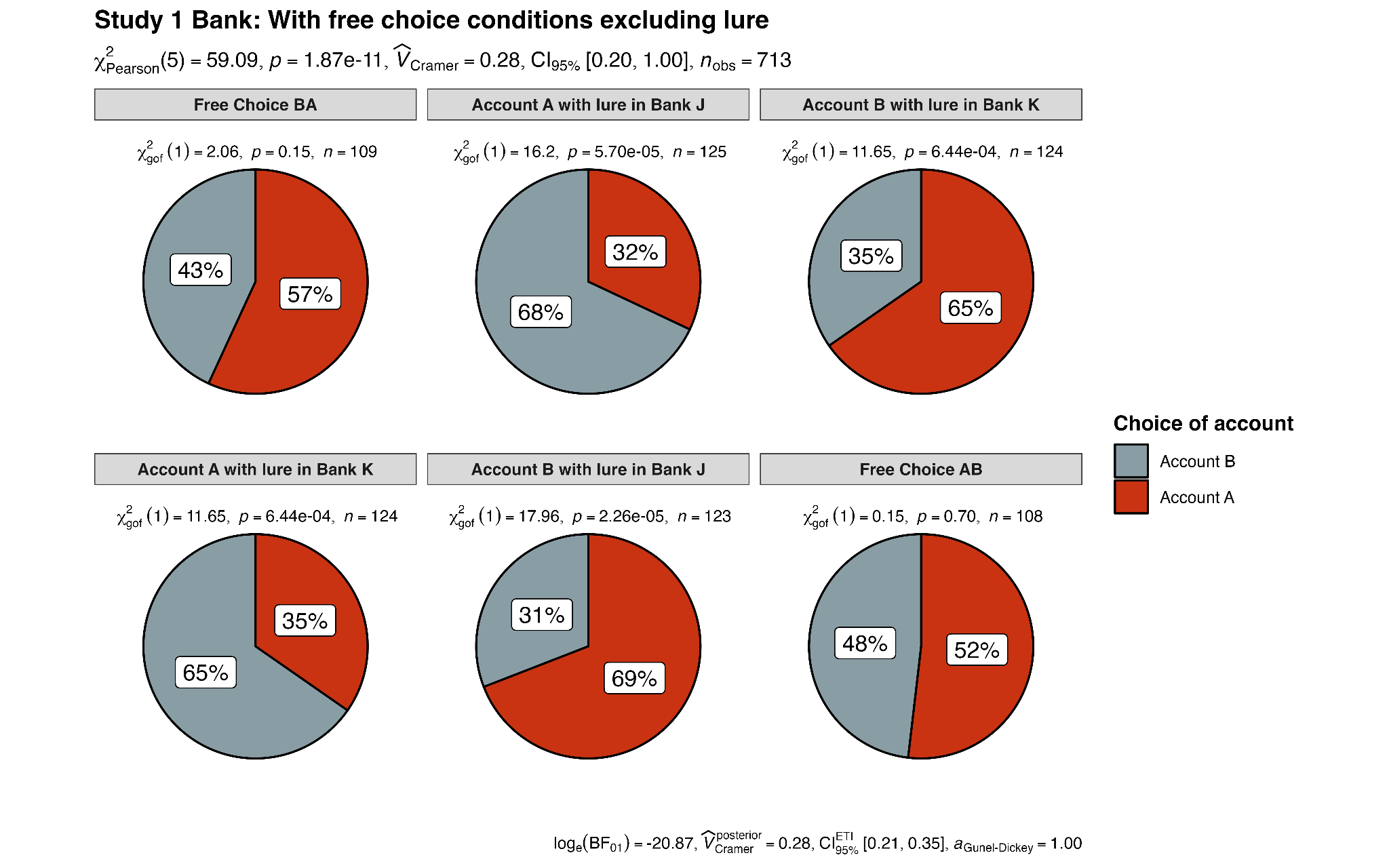
*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 8 *Study 1 Bank scenario: Comparison of conditions (excluding lure)*



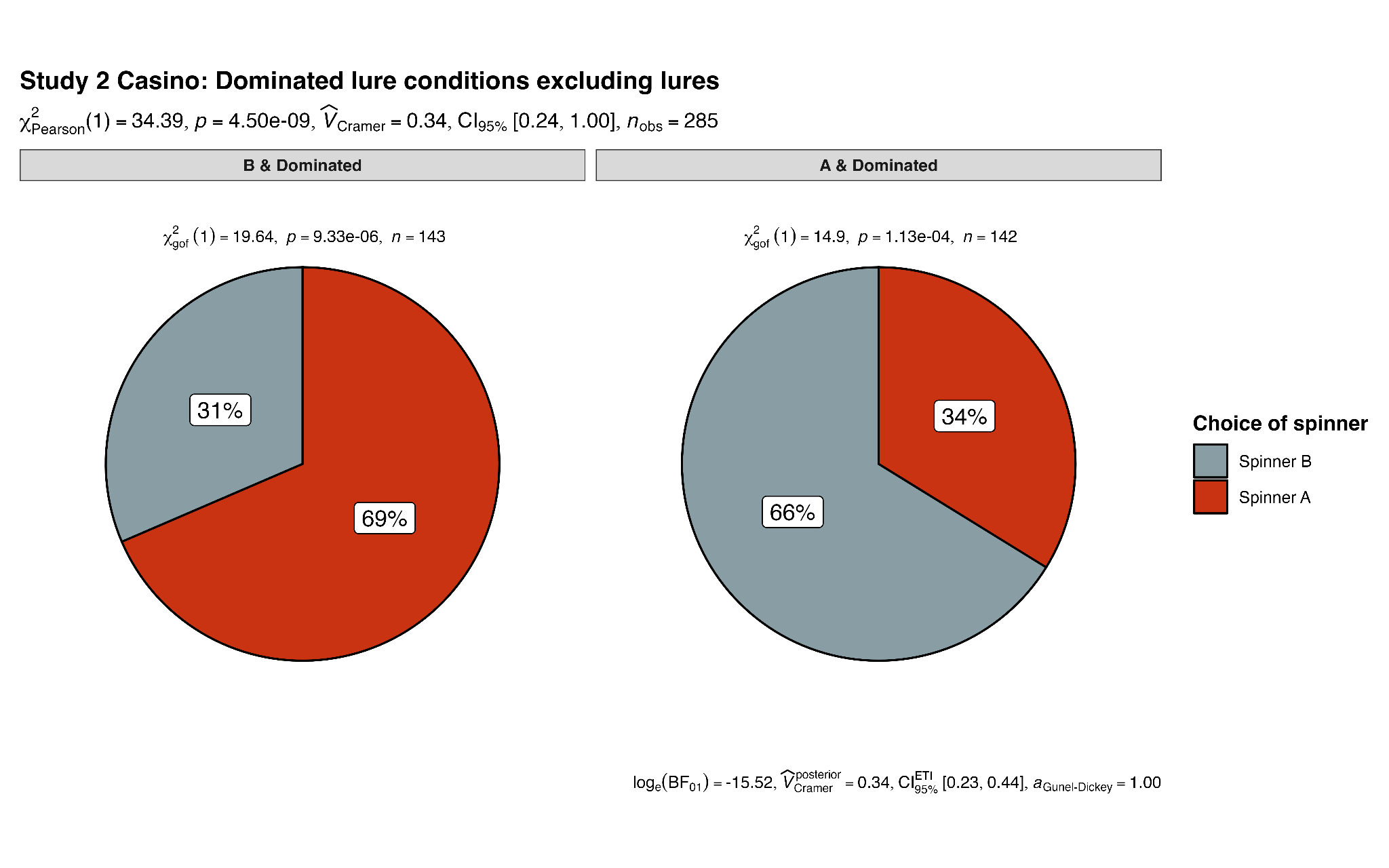
*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 9 *Study 1 Bank scenario: Comparison between replication and no-grouping conditions (excluding lure)*



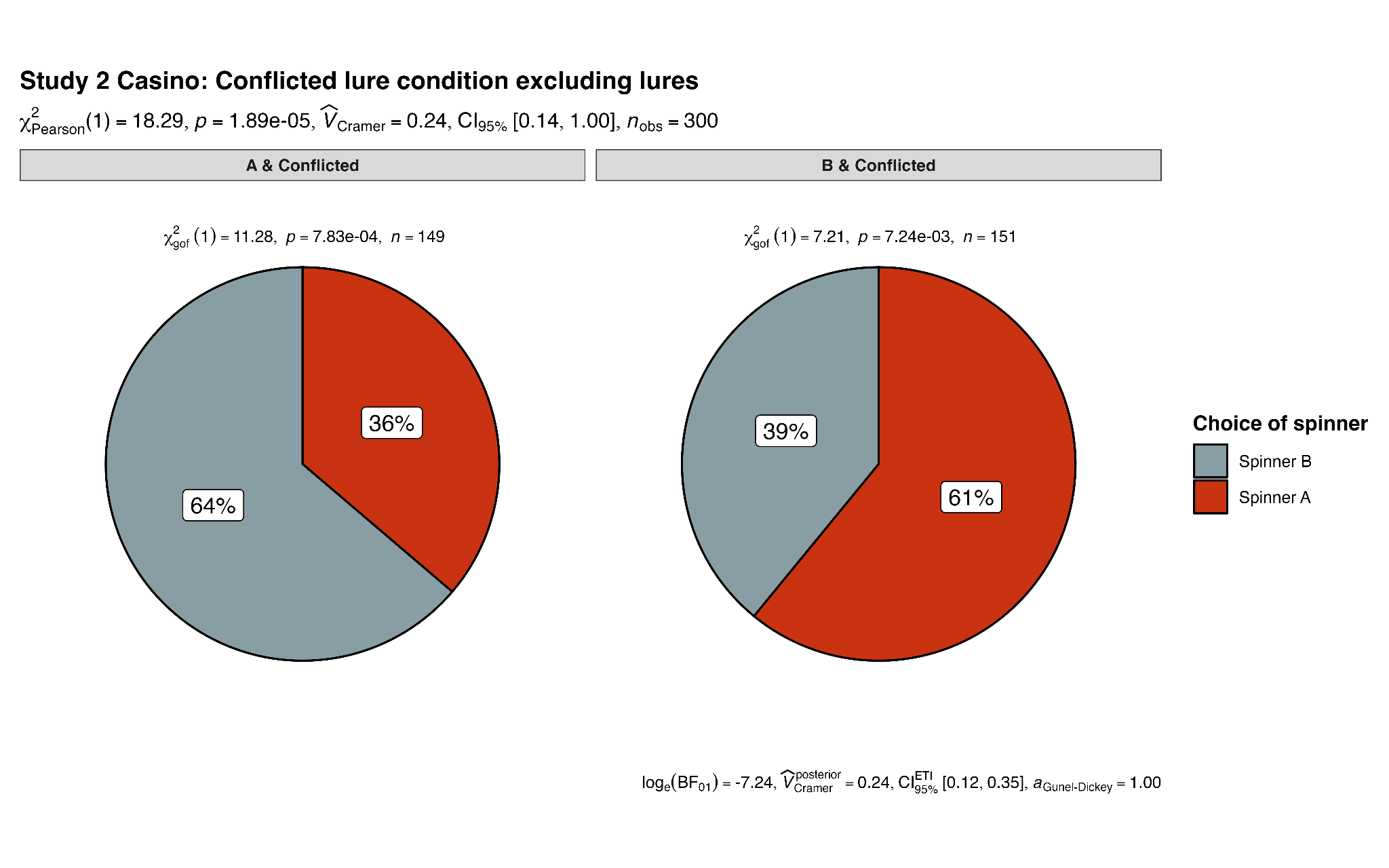
*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 10 *Study 2 Casino scenario: Comparison of spinner selection between dominated lure conditions (excluding lure)*



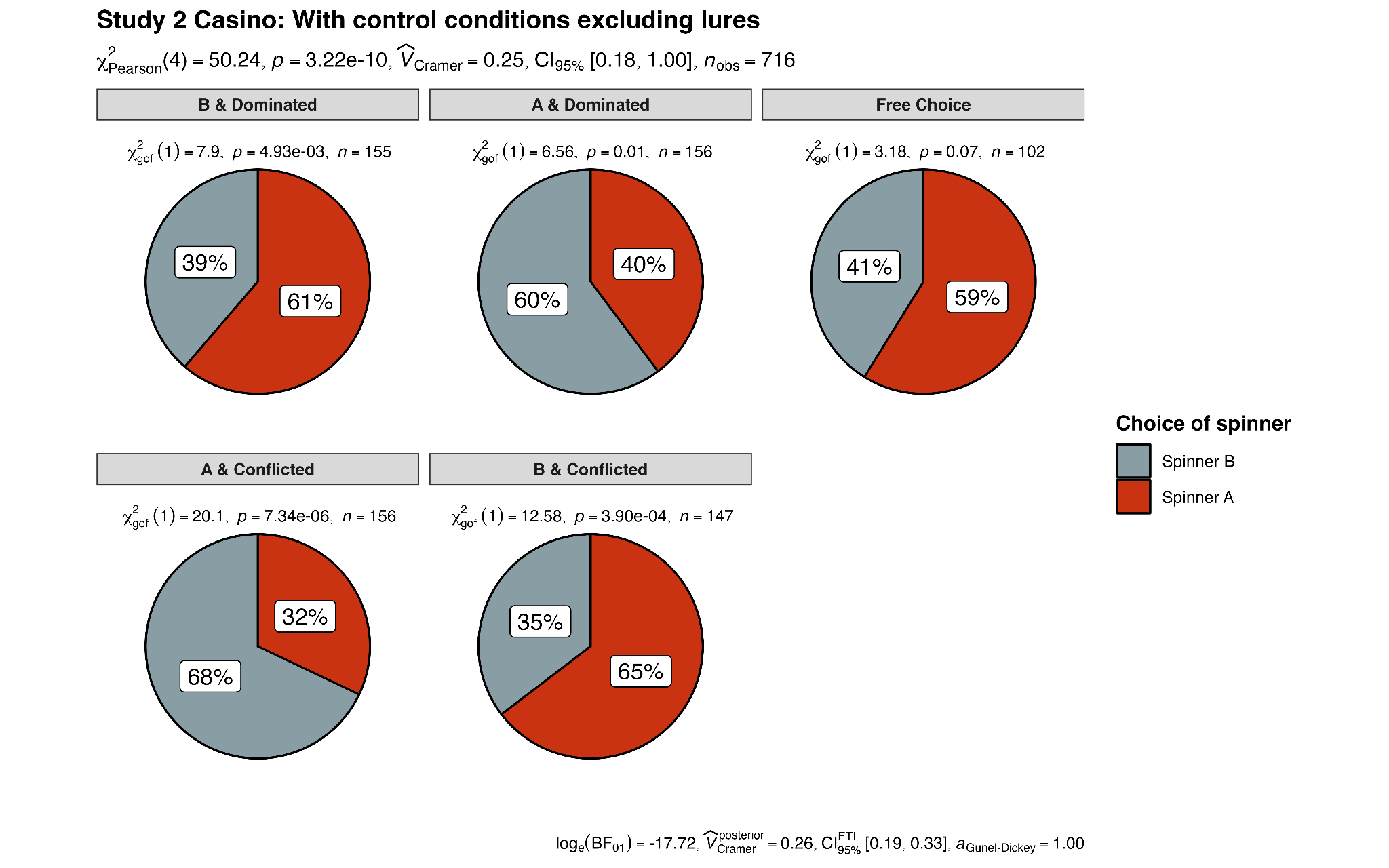
*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 11 *Study 2 Casino scenario: Comparison of spinner selection between conflicted lure conditions (excluding lure)*



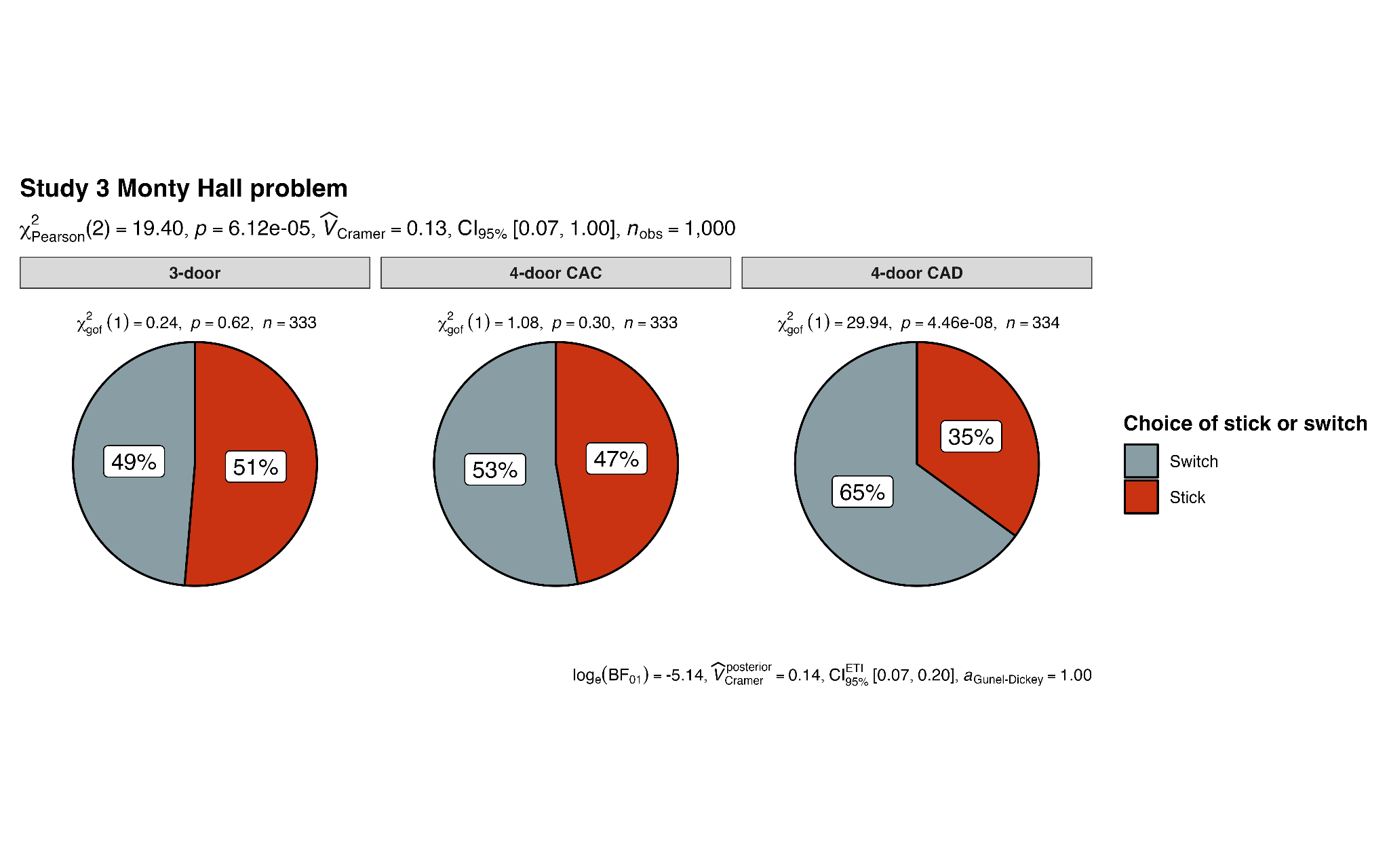
*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 12 *Study 2 Casino scenario: Comparison of spinner selection between replication and control conditions (excluding lure)*



*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

###### Figure 13 *Study 3 Monty Hall problem scenario: Comparison of conditions*



*Note*: The upper confidence interval is set to 1.00 by “ggstatsplot” (Patil, 2021) because they assume one-sided CIs. Please refer to Table 15 for confidence intervals.

## Extensions

We conducted a Fisher’s exact test and a chi-square test for each scenario along with the neutral control conditions after excluding choices of lures.

For the nightclub scenario, we found support for differences in the number of participants choosing each target nightclub between the condition the “Cherish North” condition (122 out of 166, 73%), “Cherish South” condition (120 out of 166, 72%), “Diesel North” condition (135 out of 167, 81%), “Diesel South” condition (120 out of 167, 72%), “No-grouping CD” condition (118 out of 167, 71%) and “No-grouping DC” condition (103 out of 167, 62%) after excluding lures X2(5) = 51.61, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .27, 95% CI [0.20, 0.33]. Among all posthoc comparisons involving the neutral control conditions, more participants chose Club Diesel in the “No-grouping CD” condition (66 out of 167, 40%) than in the “Diesel North” condition (47 out of 167, 28%; adj. *chisq* *p* = .003, adj. *fisher’s p* = .002) and the “Diesel South” condition (42 out of 167, 25%; adj. *chisq* *p* = .004, adj. *fisher’s p* = .004). Also, more participants chose Club Cherish in the “No-grouping DC” condition (54 out of 167, 32%) than in the “Cherish South” condition (34 out of 166, 20%; adj. *chisq* *p* = .001, adj. *fisher’s p* < .001). We provided a summary plot in Figure 9.

For the bank scenario, we found support for differences in the number of participants choosing each target account between “Account A with lure in Bank J” condition (125 out of 167, 75%), “Account B in Bank J” condition (123 out of 166, 74%), “Account A in Bank K” condition (124 out of 167, 74%), “Account B in Bank K” condition (124 out of 167, 74%), “No-grouping AB” condition (108 out of 166, 65%), and “No-grouping BA” condition (109 out of 167, 65%) after excluding lures X2(5) = 59.09, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .29, 95% CI [0.22, 0.35]. Among all posthoc comparisons involving the neutral control conditions, more participants chose Account A in the “No-grouping BA” condition (62 out of 167, 37%) than in the “Account A with lure in Bank K” condition (43 out of 167, 26%; adj. *chisq* *p* = .003, adj. *fisher’s p* = .002) and the “Account A with lure in Bank J” condition (40 out of 167, 24%; adj. *chisq* *p* < .001, adj. *fisher’s p* < .001). Also, more participants chose Account A in the “No-grouping AB” condition (56 out of 166, 34%) than in the “Account A with lure in Bank K” condition (43 out of 167, 26%; adj. *chisq* *p* = .020, adj. *fisher’s p* = .019). We provided a summary plot in Figure 7.

For the casino scenario, we found support for differences in the number of participants choosing each target spinner between the condition the “A and Dominated” condition (142 out of 199, 71%) and “B and Dominated” condition (143 out of 200, 72%), “A and conflicted” condition (149 out of 201, 74%), B & conflicted (151 out of 200, 76%), and the “No-grouping” condition (97 out of 200, 49%) after excluding lures X2(4) = 53.05, *chisq* *p* < .001, *fisher’s p* < .001; *V* = .28, 95% CI [0.21, 0.35]. Among all posthoc comparisons involving the neutral control condition, more participants chose Spinner A in the “No-grouping” condition (49 out of 200, 25%) than in the “A and Dominated” condition (48 out of 199, 24%; adj. *chisq* *p* = .024, adj. *fisher’s p* = .019). However, more participants chose Spinner A in the “B and Dominated” condition (98 out of 200, 49%) than in the “No-grouping” condition (adj. *chisq* *p* = .015, adj. *fisher’s p* = .014). We provided a summary plot in Figure 12.

[The following sections are reserved for case of failing to conclude a successful replication]

## Confirmatory analyses in case of failing to find support for replication

### 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-registered that if we failed to find support for our hypotheses, we would examine order as a moderator (without outlier exclusions), and if there is an indication for order as impacting results, we will report rerunning the analyses for the failed study by focusing on the participants that completed each study when it is displayed first. To compensate for multiple comparisons and the increased likelihood of capitalising on chance, we set the alpha for the additional analyses to a stricter .005.

[TBD conclusion based on our experience with a unified design so far: We found no differences in conclusions]

### Quantifying support for the null: Bayesian analyses

We pre-registered that in case we failed to find support for the hypothesis for any of the studies, we would run a complementary Bayesian analysis for that study using a prior of 0.707 to quantify support for the null.

# Discussion

[To be completed in Stage 2 following data collection]

## Limitations and future directions

[Planned discussion in Stage 2: Following on Dr./Prof. Hu Chuan-Peng’s comment, we will discuss generalizability of the results to other populations]

[Planned discussion in Stage 2: Following on Dr./Prof. Gakuto Chiba’s comment, we will discuss the validity of online surveys and its reproducibility and robustness of the findings of the target article.]

# Conclusion

[To be completed in Stage 2 following data collection]

# References

Bastardi, A., & Shafir, E. (1998). On the Pursuit and Misuse of Useless Information. *Journal of Personality and Social Psychology*, *75*(1), 19–32. <https://doi.org/10.1037/0022-3514.75.1.19>

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57, 289--300. <http://www.jstor.org/stable/2346101>.

Bown, N. J. (2006). *The lure of choice: The impact of increased choice on decision making* (Doctoral dissertation, University of Leeds). Retrieved from <https://etheses.whiterose.ac.uk/26110/1/485960.pdf>

Bown et al. (2003). The lure of choice. <https://doi.org/10.1002/bdm.447>

Brenner, L., Rottenstreich, Y., & Sood, S. (1999). Comparison, Grouping, and Preference. *Psychological Science*, *10*(3), 225–229. <https://doi.org/10.1111/1467-9280.00141>

Bryant, L. D., Bown, N., Bekker, H. L., & House, A. (2007). The lure of “patient choice.” *British Journal of General Practice*, *57*(543), 822–826.

Bulus M (2023). *pwrss: Statistical Power and Sample Size Calculation Tools*. R package version 0.3.1, [https://CRAN.R-project.org/package=pwrss](https://cran.r-project.org/package=pwrss).

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](https://psycnet.apa.org/doi/10.3758/BF03193146)

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

Hilton, J.F. (2005). Yates's Continuity Correction. In Encyclopedia of Biostatistics (eds P. Armitage and T. Colton). <https://doi.org/10.1002/0470011815.b2a10062>

Hsee, C. K. (1996). The Evaluability Hypothesis: An Explanation for Preference Reversals between Joint and Separate Evaluations of Alternatives. *Organizational Behavior and Human Decision Processes*, *67*(3), 247–257. <https://doi.org/10.1006/obhd.1996.0077>

Huber, J., Payne, J. W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *The Journal of Consumer Research*, *9*(1), 90–98. <https://doi.org/10.1086/208899>

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>

Kahneman, D., & Ritov, I. (1994). Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method. *Journal of Risk and Uncertainty*, *9*(1), 5–37. <https://doi.org/10.1007/BF01073401>

Kassambara A (2023). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. R package version 0.7.2, <https://rpkgs.datanovia.com/rstatix/>.

Klein, N., & O’Brien, E. (2018). People use less information than they think to make up their minds. *Proceedings of the National Academy of Sciences - PNAS*, *115*(52), 13222–13227. <https://doi.org/10.1073/pnas.1805327115>

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*. <https://doi.org/10.15626/MP.2018.843>

Mangiafico SS (2023). *rcompanion: Functions to Support Extension Education Program Evaluation*. Rutgers Cooperative Extension, New Brunswick, New Jersey. version 2.4.34, [https://CRAN.R-project.org/package=rcompanion/](https://cran.r-project.org/package=rcompanion/).

Nowlis, S. M., & Simonson, I. (1997). Attribute-Task Compatibility as a Determinant of Consumer Preference Reversals. *Journal of Marketing Research*, *34*(2), 205–218. <https://doi.org/10.1177/002224379703400202>

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>

Olsson-Collentine, A., van Assen, M. A. L. M., & Wicherts, J. M. (2020). Heterogeneity in direct replications in psychology and its association with effect size. *Psychological Bulletin*. <https://doi.org/10.1037/bul0000294>

Patil, I. (2021). Visualisations with statistical details: The'ggstatsplot'approach. *Journal of Open Source Software*, *6*(61), 3167. <https://doi.org/10.21105/joss.03167>

Peters GY, Gruijters S (2023). *ufs: A collection of utilities*. R package version 0.5.5, [https://ufs.opens.science](https://ufs.opens.science/).

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 (2022). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from <https://cran.r-project.org>. (R packages retrieved from CRAN snapshot 2023-04-07).

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice Bracketing. *Journal of Risk and Uncertainty*, *19*(1/3), 171–197. <https://doi.org/10.1023/A:1007879411489>

Selker, Ravi, Jonathon Love, and Damian Dropmann. 2023. *jmv: The “jamovi” Analyses*. [https://CRAN.R-project.org/package=jmv](https://cran.r-project.org/package=jmv).

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1-2), 11-36. <https://doi.org/10.1016/0010-0277(93)90034-s>

Sharfman, M. P., & Dean Jr, J. W. (1997). Flexibility in Strategic Decision Making: Informational and Ideological Perspectives. *Journal of Management Studies*, *34*(2), 191–217. <https://doi.org/10.1111/1467-6486.00048>

Siddiqi, H. (2009). Is the lure of choice reflected in market prices? Experimental evidence based on the 4-door Monty Hall problem. *Journal of Economic Psychology*, *30*(2), 203–215. <https://doi.org/10.1016/j.joep.2008.07.004>

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological science*, 26(5), 559-569. <https://doi.org/10.1177/095679761456734>

Sood, S., Rottenstreich, Y., & Brenner, L. (2004). On Decisions That Lead to Decisions: Direct and Derived Evaluations of Preference. *The Journal of Consumer Research*, *31*(1), 17–25. <https://doi.org/10.1086/383420>

Suzuki, S. (1997). Effects of number of alternatives on choice in humans. *Behavioural Processes*, *39*(2), 205–214. <https://doi.org/10.1016/S0376-6357(96)00049-6>

Suzuki, S. (2000). Choice between Single-Response and Multichoice Tasks in Humans. *The Psychological Record*, *50*(1), 105–115. <https://doi.org/10.1007/BF03395345>

The JAMOVI project (2023). JAMOVI. (Version 2.4) [Computer Software]. Retrieved from <https://www.jamovi.org>.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the rationality of choice. *Science*, *211*(4481), 453-458. <https://doi.org/10.1126/science.7455683>

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management science*, *39*(10), 1179-1189. <https://doi.org/10.1287/mnsc.39.10.1179>

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). <https://doi.org/10.1525/collabra.77859>

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, **4**(43), 1686. [doi:10.21105/joss.01686](https://doi.org/10.21105/joss.01686).

Yeung, S. K., Yay, T., & Feldman, G. (2022). Action and inaction in moral judgments and decisions: Meta-analysis of omission bias omission-commission asymmetries. *Personality and Social Psychology Bulletin*, *48*(10), 1499-1515. <https://doi.org/10.1177/01461672211042315>

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>

1. We reached out to the original authors and they shared with us that Bown et al. (2003) summarised four experiments out of the lead author’s doctoral thesis - Bown’s (2006), which offers a much more detailed description of those and other unpublished studies. We are grateful for their responsiveness and support. [↑](#footnote-ref-2)