**Do pain and effort increase prosocial contributions?:   
Revisiting the Martyrdom Effect with a   
replication and extensions Registered Report of Olivola and Shafir (2013)  
[Stage 1]**

Yim Tung (Emanuel) Cheng  
ORCID: <https://orcid.org/0009-0004-4233-0670>  
OSF: <https://osf.io/vpurx/>  
Department of Psychology, The University of Hong Kong  
[emanuel@connect.hku.hk](mailto:emanuel@connect.hku.hk) /[emanuelchengstudy@gmail.com](mailto:emanuelchengstudy@gmail.com)

^Gilad Feldman  
ORCID: 0000-0003-2812-6599  
OSF: <https://osf.io/m4kfe/>   
Department of Psychology, University of Hong Kong  
[gfeldman@hku.hk](mailto:gfeldman@hku.hk) / [giladfel@gmail.com](mailto:giladfel@gmail.com)

^Corresponding author

## Author bios:

Yim Tung (Emanuel) Cheng 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 judgment and decision-making.

## Declaration of Conflict of Interest:

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

## Financial disclosure/funding:

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

## Authorship declaration:

Yim Tung (Emanuel) CHENG 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 license is applied to the AAM arising from this submission. ([clarification](https://bit.ly/rrs-primer))

## Citation of the target research article:

Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of behavioral decision making*, *26*(1), 91-105.

## Acknowledgements (with ORCID):

## Contributor Roles Taxonomy

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

# 

# PCIRR-Study Design Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Question | Hypotheses | Sampling plan | Analysis plan | Statistical tests rationale | Interpretation | Theory that could be shown wrong by the outcomes |
| Study 4:  How are willingness to participate, donate, and perceived meaning impacted by effort/pain involved in fundraising activity (painful-effortful vs. easy-enjoyable)? | People donate more to and perceive higher meaningfulness in fundraisers that involve painful-effortful activities, compared to fundraisers that involve easy-enjoyable activities. | 1350 US Americans recruited online through Prolific | Pearson's correlation analysis, Chi-Square test, ANOVA | This is a replication. We follow the statistical analyses of the target article. | We examine the  replicability of  Olivola & Shafir (2013)  based on the  replication  comparison criteria  by Lebel et al.  (2019).We assess replication based on signal, direction, and when possible - overlap of confidence intervals. | Martyrdom Effect |
| Study 5:  How do cause (joy versus pain) and fundraising (effortful vs. easy) interact in impacting willingness to participate and donation amount? | People are more willing to participate in fundraisers that involve painful-effortful activities when the cause is associated with human suffering.  People are more willing to participate in fundraisers that involve easy-enjoyable events when the case is human enjoyment | Martyrdom Effect moderated by cause |
| Study 3:  How are donations associated with participation pain/effort, compared to with regular pricing | The association between run distance and price is stronger than the association between distance and donation to the charity run. | Correlations, and comparison of correlations | Attribute substitution strategy theory versus Martyrdom Effect |

# Abstract

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

The Martyrdom Effect is the phenomenon that people donate more to charity when it involves personal effort or suffering. In a Registered Report experiment with a US American sample on Prolific (*N* = 1350), we conducted a replication of Experiments 3, 4, and 5 from Olivola and Shafir (2013). [The following is a demo placeholder based on the random simulated and will be updated following data collection.] We found support for the effects of [...] (effects + 95% CI). We found mixed support for the effect of [...] (effects + 95% CI). However, we failed to find support for the effects of [...] (effects + 95% CI). Extending the replication, we [found/failed to find] support for [...]. Overall, we concluded that [...]. Materials, data, and code are available on: <https://osf.io/yu25a/>

*Keywords:* Martyrdom Effect; altruism; charitable giving; judgment and decision making; registered report; replication;

# 

# Do pain and effort increase prosocial contributions?: Revisiting the Martyrdom Effect with a replication and extensions Registered Report of Olivola and Shafir (2013)

[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

The Martyrdom Effect is the phenomenon that in charitable giving decisions with a donation matching scheme, individuals are willing to donate more when the donation involves personal effort or sacrifice. In such donation matching schemes, every donation made is matched by a donation of the same amount by another donor, effectively doubling a donor’s contribution, aiming to increase the motivation of donors to donate and donate more (e.g., Caviola & Greene, 2023). Olivola and Shafir (2013) demonstrated that individuals were more inclined to donate more in fundraisers (with a donation matching scheme) when the fundraiser activities required effort, such as running long distances or fasting, compared to more leisurely activities like attending a picnic. Their findings have been impactful and offered a new perspective on altruistic behavior and charitable giving.

We conducted a replication and extension Registered Report of Olivola and Shafir (2013) Studies 3, 4, and 5, with several goals. Our first goal was to conduct an independent, close, comprehensive, and well-powered replication of the main three studies from a classic article on the Martyrdom Effect. Our second goal was to improve on the target article’s design and methodology with adjustments and extensions that would help address open remaining questions and provide additional new insights.

We begin by introducing the literature on the Martyrdom Effect and the chosen article for replication - Olivola and Shafir (2013). We then discuss our motivations for the current replication and review Olivola and Shafir (2013) as our chosen article for replication. We then outline our chosen studies from the target article, their experimental design, and our adaptations and extensions.

## Martyrdom Effect

The Martyrdom Effect describes individuals enduring suffering for a cause they believe in and care about. This effect is named after the concept of martyrdom, where individuals make a sacrifice, often involving pain and suffering, to support their faith, beliefs, or causes. Analogously, in the context of charitable giving, people ascribe more meaning and value to charitable actions that involve pain or effort, thereby increasing their willingness to contribute towards that cause. For example, Olivola and Shafir (2013) showed that someone might be more inclined to donate to a charity when the fundraising activity involves a challenging task, such as a long-distance run or fasting, compared to an easy and enjoyable event​​ like a picnic. Xygalatas et al. (2013) demonstrated a related effect showing that those taking part in high-ordeal rituals made more prosocial donations compared to those who participated in low-ordeal rituals.

The argument was that anticipated pain and effort lead to greater prosocial contributions by the process of ascribing greater meaning to the experience of contributing and therefore to the contribution itself. This relates to theories like cognitive dissonance (Festinger, 1957) and self-perception theory (Bem, 1967), which focus on the overcoming of obstacles as enhancing the value of a goal. The Martyrdom Effect differs from those in that the mere prospect of pain or effort, rather than the actual experience, enhances the willingness to engage in that behavior.

Olivola and Shafir (2013) also argued that the Martyrdom Effect is not simply a cue for value, as suggested by attribute substitution strategy (Kahneman & Frederick, 2002). The concept of attribute substitution refers to a cognitive process in which individuals substitute a complex, less accessible attribute, like the value of a donation, with a more readily available and simpler attribute, like the amount of pain or effort. This would suggest that the reason why people donate more in more effortful tasks is because level of effort serves as a proxy for value (importance, impact, etc.). Olivola and Shafir (2013) aimed to contrast the two explanations in their Experiment 3, examining whether it is the mere presence of effort or the extent of the effort involved. They experimentally varied the hypothetical running distances, with one group of participants indicating how much they would donate to run that distance, and another group of participants indicating how much money they would ask for to run that distance (unrelated to charity). They found no support for an association between running distance and donations, yet support for an association between running distance and pricing. They concluded this as evidence of the “Martyrdom Effect” as different from the attribute substitution explanation and other theories arguing for a U-shaped relationship between pain–effort and behavior (e.g., goal setting; Locke & Latham, 2006; Trope & Fishbach, 2000). We note that the target article framed the effect as having no association for donations (null effect) and having an association for pricing, to avoid testing the null with Null Hypothesis Significance Testing, we reframing this to a joint hypothesis of a comparison of the two associations, predicting a stronger association for pricing than for donations.

## Choice of study for replication: Olivola and Shafir (2013)

We embarked on a replication and extension Registered Report of Olivola and Shafir (2013). We aimed to revisit the phenomenon to examine the reproducibility and replicability of the findings with an independent pre-registered well-powered close replication. 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 Olivola and Shafir (2013) based on several factors: its academic and practical impact, the potential for improvements in theory and methodology, and the absence of direct replications.

Olivola and Shafir (2013) had an impact on the literature of altruism and philanthropy in the domains of social psychology and judgment and decision-making. As of January 2024, it has amassed over 300 citations on Google Scholar. This work has catalyzed a series of important theoretical and empirical follow-ups, such as work by Inzlicht et al. (2018) on the complex nature of effort traditionally mostly thought of as a cost, yet, phenomena such as the Martyrdom Effect showing that it can contribute positively to perceived value of an outcome. This line of research also has implications for the psychology of (in)effective altruism (Berman et al., 2018; Caviola et al., 2020; Caviola et al., 2021; Chan & Feldman, 2024), that aims to examine and better understand the psychological processes related to the growing movement of Effective Altruism (MacAskill, 2015).

We saw much potential in clarifying and extending both theory and methodology in Olivola and Shafir (2013). Studies 4 and 5 had very similar designs contrasting effortful and easy looking, yet employed different effortful fundraisers (Study 4: run; Study 5: fasting), for different causes (Study 4: victims of war and genocide; Study 5: starving children and public park), and examined their impact on slightly different dependent variables (Study 4: meaning). The two studies employed different methodological decisions (Study 5: removal of outliers using 3 IQR rule, and of those having difficulties reading English, were distracted, gave incoherent responses, or participated in a different study). They also had slightly different findings, as - for example - in Study 5 there was no support for willingness to participate, whereas in Study 4 there was higher willingness to participate in the easy fundraiser, as well as a large main effect of effort on donation in Study 4 (*d* = .61) yet with an interaction in Study 5 with main effects per each cause condition just around the alpha threshold (*p* = .062; *p* < .04). Given the comprehensive multiple studies and the different designs, it is challenging to conclude the overall effects and the takeaway insights of the target article, and to our knowledge, there are currently no published direct replications of their studies that aimed to revisit and clarify these findings. With a well-power replication integrating Studies 4 and 5 into a unified design, we can contrast the different fundraisers against the different causes and examine all suggested dependent variables to provide a more comprehensive test of the theory. By adding a dependent variable of perceived impact and manipulation checks of perceived effort and cause importance missing in the target article, we can begin to understand the differences between the two studies, and look at more continuous categories of effortful and cause importance.

## Chosen studies for replication: Studies 3, 4, and 5

We focused our investigation on the studies examining the willingness to donate and donation amount in relation to the pain and effort involved in fundraising: Studies 3, 4, and 5.

We chose to combine Studies 4 and 5 into a single unified design for a more direct and comprehensive investigation contrasting the two types of painful-effortful fundraisers used in each of the studies compared to the one easy event that was used by both studies. We also manipulated the three types of causes used separately in the two studies, examining joint impact on both willingness to participate, donation amount, perceived meaningfulness, perceived impact, and manipulation checks of perceived effort and importance.

Study 3 was designed to address an alternative explanation to the Martyrdom Effect, in that effort (running distance) had a weaker association with donation amount compared to price (amount of money asked for participating), indicating that it is not about the level of effort, but rather effort itself. We reframed the hypotheses from a null hypothesis, and simplified and updated the methodology to allow for a clearer comparison between the two conditions, on both item (distance) and participant level.

We chose not to replicate Studies 1a and 1b as their designs were mostly covered by the designs of the more complex and informative Studies 4 and 5 (contrasting effortful versus easy activities) and our extensions (which address baseline preferences when both effortful and easy presented together). We chose not to replicate Study 2 as it involved real money and actual pain, where participants in a public goods game made larger contributions when doing so was expected to be painful. We felt that it would be better to first successfully revisit the baseline demonstration studies using hypothetical scenarios in Studies 4 and 5, before embarking on the more complex and costly replication of Study 2.

## Olivola and Shafir (2013): Hypotheses and findings

We summarized the hypotheses and findings of Studies 3, 4, and 5 in Tables 1 and 2. We aimed to stay as faithful as possible to the hypotheses in Olivola and Shafir (2013), and when those were not clearly stated - we provided an approximation based on their designs. In Study 3 the target used a null hypothesis using Null Hypothesis Significance Testing, which we reformulated by adding the alternative hypothesis, and by focusing on the contrast between the two conditions with a combined hypothesis. Similarly, in Study 5 we reformulated the hypotheses to more clearly state the main effects for each condition and then the expected interaction between the two.

Table 1  
*Olivola and Shafir (2013) Studies 3, 4, and 5: Summary of hypotheses*

|  |  |  |
| --- | --- | --- |
| **Study** | **Hypothesis** | **Description** |
| 3 | 1 (null)  1 (alternative) | In fundraisers requiring effort, level of effort is not associated with donation amount (original null hypothesis deduced from target article)  In fundraisers requiring effort, level of effort is positively associated with donation amount.  (reframed from null hypothesis) |
| 2  1 and 2 combined | In activities requiring effort, level of effort is positively associated with the price people demand to participate in that activity.  The positive association between level effort and price is stronger than the association between level of effort and donation amount. |
| 4 | 3 | People donate more to and perceive higher meaningfulness in fundraisers that involve painful-effortful activities, compared to fundraisers that involve easy-enjoyable activities. |
| 5 | 4a  4b  4a and 4b combined | People are more willing to participate in fundraisers that involve painful-effortful activities when the cause is associated with human suffering.  People are more willing to participate in fundraisers that involve easy-enjoyable events when the case is human enjoyment  Interaction between 4a and 4b. |
| 5a  5b  5a and 5b combined | People are more willing to donate to fundraisers that involve painful-effortful activities when the cause is associated with human suffering.  People are more willing to donate to fundraisers that involve easy-enjoyable events when the case is human enjoyment  Interaction between 5a and 5b. |

###### Table 2 *Summary of findings: Olivola and Shafir (2013) and our replication*

|  | **Olivola and Shafir (2013)** |  | **Replication** | | **Interpretation** |
| --- | --- | --- | --- | --- | --- |
| **Dependent Variables** | **Statistics** | **ES + CI** | **Statistics** | **ES + CI** | **(signal + direction)** |
| **Experiment 3 (*N* = 202)** | | |  |  |  |
| Donations and distance | Participant level:  r(100) = .10, *p* > .05 | *r* = .10 [-0.10, 0.29] | To be added in Stage 2, see results section for results on simulated random data. | | |
| *r* = .29 [-0.18, 0.65] |
| Item-level:  r(20) = .29, *p* > .05 |
| Price and distance | Participant level:  *r*(102) = .59, *p* < .001; | *r* = .59 [0.45, 0.70] |
| Item-level: *r*(20) = .83, *p* < .001. | *r* = .83 [0.61, 0.93] |
| **Experiment 4 (*N* = 359)** | | |  |  |  |
| Willingness to participate | χ² (1) = 1.98, *p* >.05 | *V* = 0.07 [0.00, 0.15] | To be added in Stage 2, see results section for results on simulated random data. | | |
| Donation amount  by fundraiser | *F*(1, 353) = 19.38, *p* < .001 | *η*²= 0.05 [0.02, 0.10] |
| Meaningful by fundraiser | *F*(1, 353) = 5.55, *p* < .02 | *η*²= 0.01 [0.00, 0.04] |
| Donation Amount | *t*(224.61) = 4.56, *p* <.001 | *d* = 0.61  [0.44, 0.79] |
| Meaningfulness | *t*(355) = 2.36, *p* < .02 | *d* = 0.25 [0.08, 0.42] |
| **Experiment 5 (*N* = 152)** | | |  |  |  |
| Willingness to participate | χ²(3) = 21.35, *p* < .001, | *V/ϕ* = 0.38/0.37 [0.24, 0.51] | To be added in Stage 2, see results section for results on simulated random data. | | |
| Donations by cause | *F*(1, 148) = 4.61, *p* < .04 | *η*² = .03  [0.00, 0.11] |
| Fundraiser effort x cause on donations Interaction | *F*(1, 148) = 7.12, *p* < .01 | *η*² = .05  [0.01, 0.11] |
| Donations by cause (starving children vs.. public park) | *t*(117.56) = 2.25*, p* < .03 | *d* = .41[0.04, 0.69] |
| Donation by fundraiser effort  (cause starving children) | *t*(55.94) = 2.16, *p* < .04 | *d* = .58 [0.04, 0.96] |
| Donation by fundraiser effort  (building a park) | *t*(74) = 1.89, *p* = .062 | *d* = .44  [-0.02, 0.90] |

*Note*. ES = Effect size; CI = Confidence intervals (*d*: 95% η²: 90%). Interpretation based on LeBel et al. (2019) examining signal and target’s effect size overlap with replication effect confidence intervals.

In Study 4 they showed that the fundraiser type (painful–effortful vs. easy–enjoyable) influences participants' donation amounts and perceived meaningfulness. They found higher mean donation for painful–effortful events (£17.95) compared to easy–enjoyable ones (£5.74; *t*(224.61) = 4.56, *p* < .001), and painful–effortful fundraisers were rates as more meaningful (*M* = 6.46) than fundraisers that were easy–enjoyable (*M* = 5.93; *t*(355) = 2.36, *p* < .02).

In Study 5 they found differences in participation rates across different conditions, with the highest likelihood of participation observed in the [public park + picnic] condition and the lowest in the [public park + fasting] condition, as indicated by a *χ²*(3) = 21.35, *p* < .001, *ϕ*= 0.37. Furthermore, they found support for differences in donation amounts (Kruskal–Wallis: *χ²*(3) = 10.70, *p* < .02, *ϕ*= 0.27), with the largest donations occurring in the “starving children + fasting” condition. Most importantly, they reported support for a cause-fundraiser interaction effect. There was higher willingness to donate to painful-effortful fundraisers for causes related to human suffering, and a trend towards higher donations in easy-enjoyable fundraisers for causes related to human enjoyment (cause: *F*(1, 148) = 4.61, *p* < .04, *η²* = .03; cause-fundraiser interaction *F*(1, 148) = 7.12, *p* < .009, *η²* = .05). They concluded that cause can moderate the Martyrdom Effect.

In their Study 3, they addressed an alternative explanation comparing donations to costs. They found that the association between the charity run distance and the amount donated (participant level: *r*(100) = .10, ns; item level: *r*(20) = .29, ns) was weaker than the association between run distance and the price requested for running (participant level: *r*(102) = .59, *p* < .001; item-level: *r*(20) = .83, *p* < .001)[[1]](#footnote-2). The findings were meant to show that participants recognized longer distances as requiring greater effort or pain, yet did not directly translate into increased donations with distance.

## Pre-registration and open-science

We provided all materials, data, and code on: (Enter link after IPA). 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 randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection. For the purpose of the simulation, we wrote things in past tense, but no pre-registration or data collection took place yet.]

## Power and sensitivity analyses

We conducted a power analysis and calculated effect size, and confidence intervals based on a guide by Jané et al. (2024) and R (Version 4.3.2; 385 R Core team, 2023). We focused on the differences between conditions regarding the main dependent variable - donation amount, and aimed for a 95% power with an alpha of 0.05. We used GPower 3.1.9.7 (Faul et al., 2007) for power and sensitivity tests. We summarized the calculated effects for Studies 3 to 5 in Table 2. The smallest effect size for the impact of effort on donations was *d* = .41 in Study 5, requiring 130 per condition (260 overall). In our unified design in Studies 4 and 5 we had nine conditions, and therefore 1170 overall. We added a buffer of an extra 180 participants (20 participants per condition), therefore aiming for a total sample of 1350 participants (or 150 per condition).

We conducted a series of sensitivity analyses, detailed in the “Sensitivity analyses” subsection of the supplementary materials. We aimed for each condition pair to be powered for the smallest between-subject design t-test contrast (*d* = .41), yet given that we have three cause effort conditions, two effortful conditions compared to one easy, and two human suffering cause conditions compared to one human enjoyment, the contrast (without the buffer) is 260 effortful compared to 130 easy, or if cause is collapsed then 520 effortful compared to 260 easy in the combination of the human suffering conditions. Our sensitivity analyses show that when collapsing these, we should be powered to detect much weaker effects of *d* = 0.35 and *d* = 0.25, respectively. Sensitivity analyses further show that 130 per condition is powered to detect effects of *f* = 0.16 for the 2x2 interaction in the replication of Study 5, *f* = 0.12 for the 3x3 interaction in our extension of the unified design of Studies 4 and 5, *w* = 0.22 for the chi-square test in the replication of Study 4, *w* = 0.18 for the chi-square test in the replication of Study 5, and *w* = 0.14 for the unified design of Studies 4 and 5.

[Note: We will update the sensitivity analyses in Stage 2 to the exact final number of participants after data collection]

## Participants

[To demonstrate what the results would look like after data collection we simulated a dataset of 1147 participants using Qualtrics and reported our analyses below based on that dataset. Results will later be updated in full to a sample of 1350 and the real data.]

We recruited a total of 1147 US Americans on Prolific (*Mage* = 50.1, SD = 29.3; 276 females, 302 males, 271 other, 298 did not disclose). We used Prolific’s filters to restricted the location to the US using “standard sample”, we set it to “Nationality: United States”, “Country of birth: United States”, “Minimum Approval Rate: 90, Maximum Approval Rate: 100”, “Minimum Submissions: 50, Maximum Submissions: 100000”. We summarized a comparison of the target article sample and the replication samples in Table 3.

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Olivola and Shafir (2013) | | | US American on Prolific |
|  | Study 3 | Study 4 | Study 5 |  |
| Sample size | 202 | 359 (after exclusions) | 152 (after exclusions) | 1147 | |
| Geographic origin | US Undergraduate Students | UK general public | US Shoppers of a shopping mall | US American | |
| Gender | 57% (~116) female, the rest unreported | 48% (~173) female, the rest unreported | 48% (~73) male, the rest unreported | 302 males, 276 females, 271 other, 298 did not disclose | |
| Median age (years) | Unreported | Unreported | Unreported | 51 | |
| Average age (years) | Unreported | 38.37 | 34.75 | 50.1 | |
| Standard deviation age (years) | Unreported | 6.70 | 14.73 | 29.3 | |
| Age range (years) | Unreported | 19-50 | 18-86 | 0-100 | |
| Medium (location) | University | Online via Maximiles survey service | A shopping mall in US | Computer (online) | |
| Compensation | Course credit and nominal payment | Provided, with unreported type | Provided, with unreported type | Nominal payment | |
| Year | Unreported | Unreported | Unreported | 2024 | |

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

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

### Design: Replication and extension

In the target article, Studies 3, 4, and 5 were conducted separately with independent samples. We ran all the studies together in a single unified data collection, and combined Studies 4 and 5 into a unified study design. The display of scenarios and conditions 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. See “data analysis strategy” section below.]

We summarized the experimental design of our unified design for Studies 4 and 5 of the target article in Table 4, and of Study 3 in Table 5. The design of study 3’s replication was a 2 (between; donation vs. price) by 20 (between; distance: 1 to 20 miles) study design.

The design of our unified replication design for Studies 4 and 5 was a 3 (between; painful–effortful vs. easy–enjoyable vs. painful–effortful) by 3 (between; human suffering war victims vs. human suffering starving children vs. human enjoyment picnic).

###### Table 4 *Studies 4 and 5: Unified experimental design*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **IV1:Fundraiser Type  [Between]**  **IV2: Cause [Between]** | **IV1: Painful–Effortful Fundraiser (5-mile run)**  [Study 4 replication]  “Imagine a nonprofit organization sponsoring a five-mile charity run to raise money for [cause].  To attend, you must donate a certain amount of money (greater than $0). The organization will match the donations made by every person who attends the run. In other words, two dollars will be donated for every dollar collected from attendees so that the amount of money raised will be double the amount collected. All the donations raised by this run will go towards aiding [cause].  5,000 people are expected to attend. | **IV1: Easy–Enjoyable Fundraiser (picnic)**  [Studies 4 and 5 replication]  “Imagine a nonprofit organization sponsoring an outdoor charity picnic to raise money for [cause].  To attend, you must donate a certain amount of money (greater than $0). The organization will match the donations made by every person who attends the picnic. In other words, two dollars will be donated for every dollar collected from attendees so that the amount of money raised will be double the amount collected. All the donations raised by this picnic will go towards aiding [cause].  5,000 people are expected to attend.” | **IV1: Painful–Effortful Fundraiser (30-hour fast)**  [Study 5 replication]  “Imagine a nonprofit organization sponsoring a 30-hour charity fast (going without food for 30 consecutive hours) to raise money for [cause].  To attend, you must donate a certain amount of money (greater than $0). The organization will match the donations made by every person who attends the fast. In other words, two dollars will be donated for every dollar collected from attendees so that the amount of money raised will be double the amount collected. All the donations raised by this fast will go towards aiding [cause].  5,000 people are expected to attend.” |  |
| **IV2: Human Suffering 1**  Victims of war and genocide.  [Study 4 replication]  **IV2: Human Suffering 2**  Starving children in the poorest countries.  [Study 5 replication]  **IV2: Human Enjoyment**  Build a new public park in a community.  [Study 5 replication] | (the same dependent variables across all conditions)  "Please take a moment to imagine attending the [run/picnic/fast] and knowing that, as a result, your donations have been matched and will go towards aiding [cause].”  **DV1: Willingness to participate**  “Would you attend this [five-mile charity run/charity picnic/charity fast]?”  **DV2: Donation amount**  “How much money would you donate to [run/attend the picnic/attend the fast]? I would donate $ \_\_\_\_\_\_\_\_\_\_\_\_\_. ($0-$100)”  **DV3: Meaning**  “How meaningful (to you) would this experience be?”  “How meaningful (to you) would your participation in the event be?”  “How meaningful (to you) would your contribution be?”  Scale: 1 = *Not at all meaningful*; 10 = *Very meaningful*  For analysis: Mean of the three items.  **DV4: Perceived impact**  “How much of an impact would your contribution have on this charitable cause?”  Scale: 1 = *Very little impact*; 10 = *Very large impact*  **Manipulation checks**  “In your personal view, how important is this charitable cause?”  Scale: 1 = *Not at all important*; 10 = *Very important*  “How difficult and effortful is the activity for this charitable cause for you?”  Scale: 1 = *Very easy*; 10 = *Very difficult* | | | |

*Note*. There were additional background checks shared with the other study. See “controls and checks” section.

###### Table 5 *Study 3: Replication and extension experimental design*

|  |  |
| --- | --- |
| IV1: Donation/Cost [between] | IV2: Miles run (X miles from 1 to 20) [between] |
| **IV1: Donation Condition**  Donations to run the specific distance if donation is matched upon successful completion | Scenario:  "Imagine that a non-profit organization is sponsoring a charity run in order to raise money to support research to address Alzheimer’s disease. The organization will match the donations made by every person who completes the run. In other words, two dollars will be donated for every dollar collected from runners so that the amount of money raised by this run will go towards researching Alzheimer’s disease. Please indicate below the amount of money you are willing to donate for the following charity run:"  Dependent variable: **Donation per distance**  The DV varied depending on the condition:  X miles: "How much money would you be willing to donate when participating in a **X-mile** charity run (with donation matching for successfully completed runs)? I would donate $\_\_\_\_\_\_."  Scale: $0 to $100 |
| **IV1: Price Condition**  Smallest amount of money required to run the specific distance | Dependent variable: **Smallest amount required to run the specified distance**  Imagine that you were asked to run a certain number of miles, and that you will be paid for it. Please indicate below the smallest amount of money for which you would agree to run the following distance:"  The DV varied depending on the condition:  X miles: "How much money would you require to complete a **X-mile** run?  The smallest amount I would require is $\_\_\_\_\_\_."  Scale: $0 to $100 |

*Note*. There were additional background checks shared with the other study. See “controls and checks” section.

## 

## 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/e406d7c1-09be-4fb8-9078-6c8342732296/SV_40aJ0IFqYFndixE?Q_CHL=preview> ]

We reconstructed the target’s stimuli and adjusted it to an online Qualtrics survey based on the information provided in the article. We reached out to the author of the target article and are very grateful for the materials she provided which were very helpful in our reconstruction of the studies.

Participants 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 options order being rotated (yes, no, not sure) indicating confirmation of: (1) paying close attention to details and answering subsequent questions carefully, (2) agreement to having to answer attention and comprehension checks, and (3) being a native English speaker born, raised, and currently located in the US. Failing any of the three attention questions meant that the participants did not indicate consent and therefore could not embark on the study. These were followed by writing a statement indicating that they understand and agree and terms, which participants had to write correctly in order to proceed, with as many attempts as needed. Upon completion of these steps, participants proceeded to begin the survey with Study 3 and Studies 4 and 5 combined, presented in random order.

In Study 3, participants were randomly assigned in one of two conditions: donation versus price. In the donation condition, participants indicated the amount of money they were willing to donate for completing charity runs of a specific distance, ranging from 1 to 20 miles. In the price condition, participants indicated the smallest amount of money they would require to run the specific distance again varying from 1 mile to 20 miles. In both, participants were randomly assigned to rate only one distance, from 1 to 20 miles.

In Studies 4 and 5, participants were randomly assigned to one of nine conditions: one of three types of fundraising activity (a five-mile charity run, a charity picnic, or a 30-hour charity fast) and one of three different causes (aiding victims of war and genocide, feeding starving children in the poorest countries, or building a new public park in a community).

Participants read a scenario about the assigned fundraising activity for the assigned cause, with matched donations by the organizing nonprofit, effectively doubling the contribution to the cause. The scenarios were crafted to elicit a sense of engagement and to assess participants' willingness to contribute to various causes through different types of fundraising activities.

Participants indicated their willingness to attend the event, the amount they are willing to donate, perceived meaning, perceived impact, perceived effort, and perceived importance.

Participants then answered a number of controls and general checks about their perceptions regarding the three types of fundraisers, regardless of their assigned condition.

At the end of the experiment, participants answered a number of funneling and demographic questions and were debriefed.

## Manipulations

### Study 3: Donation versus price and Distance

Participants were randomly assigned to either indicate donation amount for varying distances of a charity run or a required payment (price) to run those same distances, and randomly assigned to a distance in miles, between 1 and 20 miles.

### Studies 4 and 5 combined: Fundraiser type and cause

Participants were randomly assigned to one of three fundraiser types: charity run (painful–effortful; replication of Study 4), a charity picnic (easy–enjoyable; replication of Studies 4 and 5), and fasting for 30 hours (painful–effortful; replication of Study 5), and randomly assigned to one of three causes: aid for war victims (human suffering-related cause; replication of Study 4), support for starving children (human suffering-related cause; replication of Study 5), and funding for building a new park (human enjoyment-related cause; replication of Study 5).

## Manipulation checks (extension)

To ensure that the target article’s manipulations indeed manipulated effort and cause importance, and to examine the differences between the manipulations in Study 4 and Study 5 given the unified design, we added manipulation checks. We note that this is a deviation as Olivola and Shafir (2013) did not include manipulation checks (see Table 6). This is especially important, given differences in findings between Study 4 and Study 5 in the target article, such as in the support for an effect regarding willingness to participate, and so the manipulation checks might also help compare the effortful conditions of running and fast, to examine differences in perceived difficulty, and similarly for fundraiser type. We added a measure of importance - “In your personal view, how important is this charitable cause?” (1 = *Not at all important*; 10 = *Very important*) and difficulty - “How difficult and effortful is the activity for this charitable cause for you?” (1 = *Very easy*; 10 = *Very difficult*).

## Dependent Measures

### Studies 4 and 5 combined

#### Meaningfulness

We adopted the 10-point Likert scale used in the target article. The scale for this measure included ratings of how meaningful participants found the experience, their participation, and their donation (1 = *Not at all meaningful*, 10 = *Very meaningful*; = .XX)

#### Donation amount

We examined the amount of money participants were willing to donate. Specifically, the donation scale consisted of values ranging from $0 to $100.

### Study 3: Donation vs. Price

In the donation condition, participants indicated how much they are willing to donate to participate in the charity run of the distance in the condition they were randomly assigned to (from $0 to $100).

In the price condition, participants indicate the amount of money they would require to run the distance in the condition they were randomly assigned to , serving as an indirect measure of perceived effort or inconvenience (from $0 to $100).

We note that in both we followed the target’s outlier criteria applied on their data and restricted the range that participants can indicate for their hypothetical donations, given our previous negative experience with unrestricted range in decision-making task with online samples, and we thought this to be a better strategy than a post-hoc arbitrary outlier criteria.

## Controls and general checks

In accordance with the target article we added several questions regarding participants’ background in relation to the activities used in the survey. We first asked “Do you have any medical conditions or other physical constraints that would prevent you from participating in the following types of fundraisers?” for fundraiser events “Picnic fundraiser”, “Charity run fundraiser”, and “Fasting fundraiser” (1 = “*YES - I have a medical condition or physical constraint preventing me from participating*”; 0 = “*NO - I do NOT have a medical condition or physical constraint preventing me from participating*”). We then asked regarding whether the participants have any experience with participating in any of these events (1 = “*YES - I have previously participated in such a fundraiser*”; 0 = “*NO - I have never previously participated in such a fundraiser*”), and how often they engage in such events - “How often do you engage in the following activities in your life (in general, unrelated to charity)?” (for picnic, running, or fasting; 0 = *Never engaged in this activity*; 6 = *Engage in this activity regularly*). Finally, we asked participants about perceived effort of each of the activities, even ones they were not assigned to - “How effortful are the following activities to you? (in general, unrelated to charity)” (0 = *Not effortful at all*; 6 = *Extremely effortful*).

### Deviations

We made a few adjustments with reference to the original study design, summarized in Table 6.

###### Table 6 *Replication and extension adjustments to the target article’s methods and design*

| **H** | **Studies** | **Factor** | **Target article** | **Adjustment in current study** | **Reason for change / Justifications** |
| --- | --- | --- | --- | --- | --- |
| All | All | Study  design | Participants completed the  studies in a university (study 3), online survey (study 4), and with pen and paper in a  shopping center (study 5) | Participants completed the studies on an online survey | Lower cost, higher efficiency, more diverse population (generalizability) |
| All | All | Procedures | The three 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 | All | Sample Size | *N* = 713 (3 studies combined after exclusions) | *N* = 1350 | Based on a power analysis |
| All | All | Monetary unit | United States Dollar (USD) was used in Studies 3 and 5; British Pound (£) was used in Study 4 | All conditions use the United States Dollar (USD) as the monetary unit for donation amounts | Help standardize the donation amounts across different conditions, making the data easier to compare and analyze; make the study more relatable and understandable for participants as the participant pool in the replication study is primarily from the United States. |
| All | All | Study design (Donation Range) | No donation cap was applied in studies 3, 4, and 5. | The donation range is set from $0 to $100 | Ensuring reasonable responses based on the target’s post-hoc outlier criteria. Helps standardize the measurement across participants, enhancing comparability and consistency in the data. |
| 3,4,5 | 4,5 | Extension DV | Participants were asked to rate the meaningfulness of the experience, their participation in the event, and their contribution on a 1-10 scale. | An additional DV was introduced to assess participants' perceptions of the impact of their contribution to the charitable cause. | Capture not only the emotional or existential significance (meaningfulness) of the action but also the perceived practical effect or efficacy of the contribution; offer insights into whether participants' sense of meaningfulness correlates with their perceptions of impact. |
| 3,4,5 | 4,5 | Study  design | Study 4: 1x3 with 3 DVs.  IV was painful/effortful 5-mile vs. easy-enjoyable picnic, in context of human suffering war victims cause, examining willingness, donation amount, and meaning.  Study 5: 2x2 with 2 DVs. IV1 was painful-effortful fasting vs. easy-enjoyable picnic. IV2 was human suffering aiding starving children vs. human enjoyment building a new public park. Examining willingness and donation amount. | Combining to a single 3x3 with 3 DVs unified replication experimental design, and an additional extension DV. | More efficient data collection process, reducing the complexity and costs associated with running separate studies; Facilitates direct comparison across different types of fundraisers and causes within the same participant pool, enhancing the ability to detect nuanced effects or interactions between fundraiser types and causes. |
| All | 4,5 | Manipulation checks | No manipulation checks. | Added two manipulation checks for effort and cause importance. | Ensure that the manipulations are working as intended and in the case of a failed replication to differentiate the manipulation from the dependent measure. Allows comparison of strength of the manipulations used in Study 4 and Study 5. |
| All | All | Control and general checks | Not clearly specified in the target article, used as robustness controls in Study 4. | Four questions for three activities for exploratory analyses. |  |
| All | 4 | Attention checks | Target article included several checks in their Study 4, used to exclude participants | We did not include attention checks in Study 4 | We instead had manipulation checks and other attentiveness checks in place in the outline and consent before the study begins. |

## 

## Evaluation criteria for replication findings

We aimed to compare the replication effects against 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 2 or 3 out of 3 target studies showed a signal in the same direction as the target article, a failed replication if no studies; and mixed findings if only 1 was successful.

## Replication closeness evaluation

We provided details on the classification of the replications using the criteria by LeBel et al. (2018) criteria in Table 7 below. We summarized the replication as a "close” replication.

###### Table 7 *Classification of the replication, based on LeBel et al. (2018)*

|  |  |  |
| --- | --- | --- |
| **Design facet** | **Replication** | **Details of deviation and severity** |
| Effect/hypothesis | Same |  |
| IV construct | Same |  |
| DV construct | Same |  |
| IV operationalization | Same |  |
| DV operationalization | Same |  |
| IV stimuli | Same |  |
| DV stimuli | Similar | We adjusted the donation amount range and scenarios (see deviations table). |
| Procedural details | Different | We ran all scenarios in a unified single data collection with randomized scenario presentation order. |
| Physical settings | Different | All experiments were conducted online. |
| Contextual variables | Unknown | Participants were recruited online using Prolific |
| Population (e.g., age) | Similar | We recruited participants online with a more diverse population. |
| 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

We provided an accompanying Rmarkdown on the OSF, with outputs based on simulated random data. The analyses were performed with a combination of R (Version 4.3.2) with assistance from exported R syntax from JAMOVI (Version 2.2.2). We wrote our planned analysis code using packages “ggstatsplot” (Version 0.12.2), “grateful” (Version 0.2.4), “grid” (Version 4.3.2), “jmv” (Version 2.4.11), “MBESS” (Version 4.9.3), “psych” (Version 2.3.12).

### Studies 4 and 5 combined

We conducted a comprehensive 3x3 between-group analysis, which covers all the analyses reported in the original studies, yet allows for additional insights. This design encompassed three fundraiser types (charity run, charity picnic, fasting) and three activity causes (war victims, starving children, new park).

We conducted ANOVAs on the continuous measures and chi-squares on the dichotomous measures to mirror the replications in the studies, as part of a larger 3x3 ANOVA of all conditions with post-hoc contrasts.

We also added a mediation analysis with bootstraps (replacing the older approach), to imitate the target’s analyses, though we note that we see limited value in such analysis given that meaningfulness and willingness and donation are all dependent variables in the same condition, limiting us from concluding causality.

### Outliers and exclusions

In this study, we would not pre-register classifying outliers. We included all the data collected in our analysis for those who successfully completed the entire study. We added measures aimed to conduct additional exploratory analyses in case of a replication failure.

### Order effects

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

We therefore pre-register that if we fail to find support for the core hypotheses of the target article that we rerun exploratory analyses for the failed study by focusing on the participants that completed that study first, and examine order as a moderator (without outlier exclusions). To compensate for multiple comparisons and increased likelihood of capitalizing on chance, we will set the alpha for the additional analyses to a stricter .005.

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

# Results

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

We summarized the statistical tests and effects for all studies in Table 2 with a comparison against the target article’s, with descriptives for Studies 4 and 5 in Table 8 and for Study 3 in Table 9.

## 

###### Table 8 *Studies 4 and 5 combined: Descriptives*

|  |  |  |  |
| --- | --- | --- | --- |
| IV1: Fundraiser Type (between)  IV2: Cause (between) | **5-mile run** | **Picnic** | **30-hour fast** |
| **Victims of war and genocide.** | *n* = 127 | *n* = 127 | *n* = 128 |
| Willingness to participate | 48.03% | 48.82% | 54.69% |
| Donation amount | $25.70 (30.10) | $25.58 (30.26) | $27.62 (29.65) |
| Meaningfulness | 5.42 (1.50) | 5.59 (1.70) | 5.37 (1.59) |
| Perceived impact | 5.74 (2.87) | 5.77 (2.92) | 5.74 (2.91) |
| **Starving children in the poorest countries.** | *n* = 127 | *n* = 128 | *n* = 128 |
| Willingness to participate | 43.31% | 55.47% | 40.94% |
| Donation amount | $19.64 (26.49) | $26.70 (27.65) | $28.89 (27.52) |
| Meaningfulness | 5.53 (1.57) | 5.31 (1.48) | 5.19 (1.54) |
| Perceived impact | 5.66 (2.96) | 5.44 (2.75) | 5.51 (2.80) |
| **Build a new public park in a community.** | *n* = 128 | *n* = 127 | *n* = 127 |
| Willingness to participate | 50.78% | 51.97% | 47.24% |
| Donation amount | $24.08 (26.23) | $27.62 (28.58) | $22.28 (27.17) |
| Meaningfulness | 5.62 (1.64) | 5.45 (1.56) | 5.62 (1.68) |
| Perceived impact | 5.77 (2.71) | 5.54 (2.90) | 5.87 (2.84) |

*Note:* Format = mean (standard deviations);n = indicates sample size; % = percentage

#### 

## Replication of Study 4

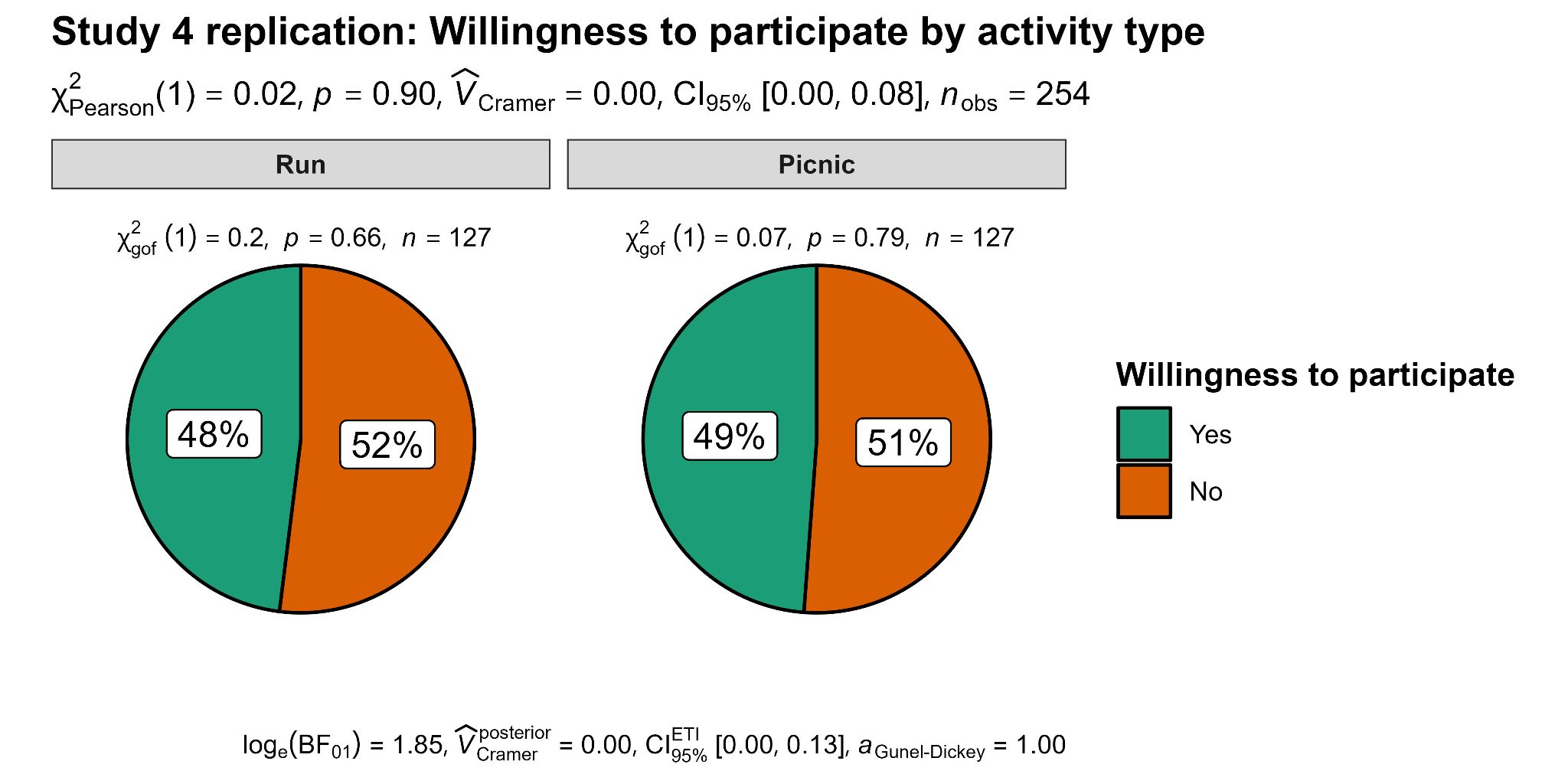
Mirroring the target article’s Study 4, in the unified Studies 4 and 5 design we focused on war victims cause contrasting fundraiser conditions run and picnic.

### Willingness to participate

We conducted a chi-square test (two-tail) and found no support for differences in willingness to participate between the effortful (run) condition (*n* = 127; Yes = 48.0%) and the easy (picnic) condition (*n* = 127; χ²(1) = 0.02, *p* = 0.90; Cramer's V = 0.00, 95% CI [0.00, 1.00]). We provided a summary plot in Figure 1.

###### 

###### Figure 1 *Study 4 (replication) - willingness to participate by fundraiser type*



*Note*. Created using ggstatsplot (Patil, 2021)

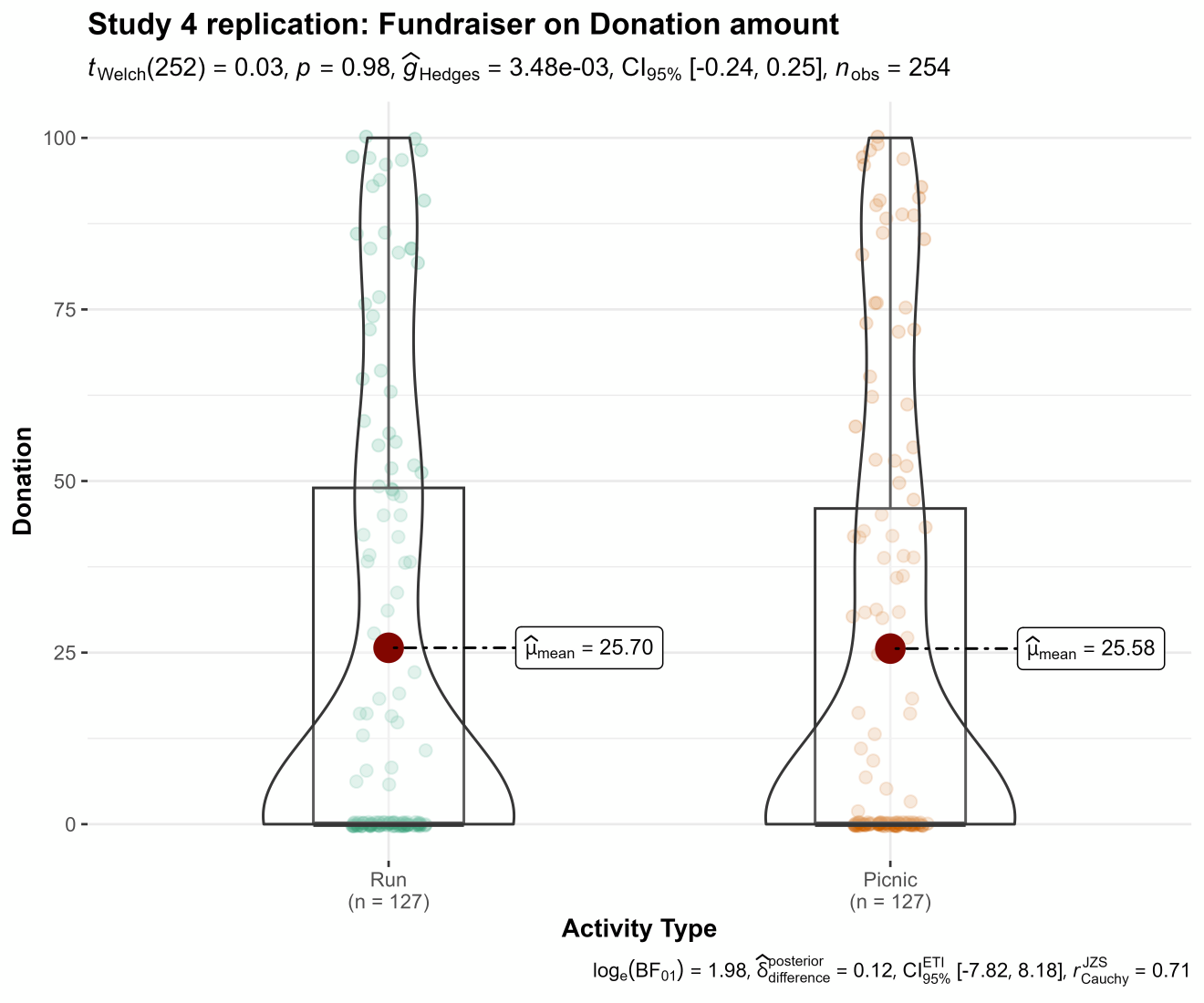
### 

### Donation amount

We conducted an independent samples t-test and found no support for differences in donation amount between the effortful (run) condition (*n* = 127, *M* = 25.7, *SD* = 33.9) and the easy (picnic) condition (*n* = 127, *M* = 25.6, *SD* = 33.7; *t*(252) = 0.03, *p* = .98; *d* = .00, 95% CI [-0.24, 0.25]). We provided a summary plot in Figure 2.

###### 

###### Figure 2 *Study 4 replication - donation amount by fundraiser type*

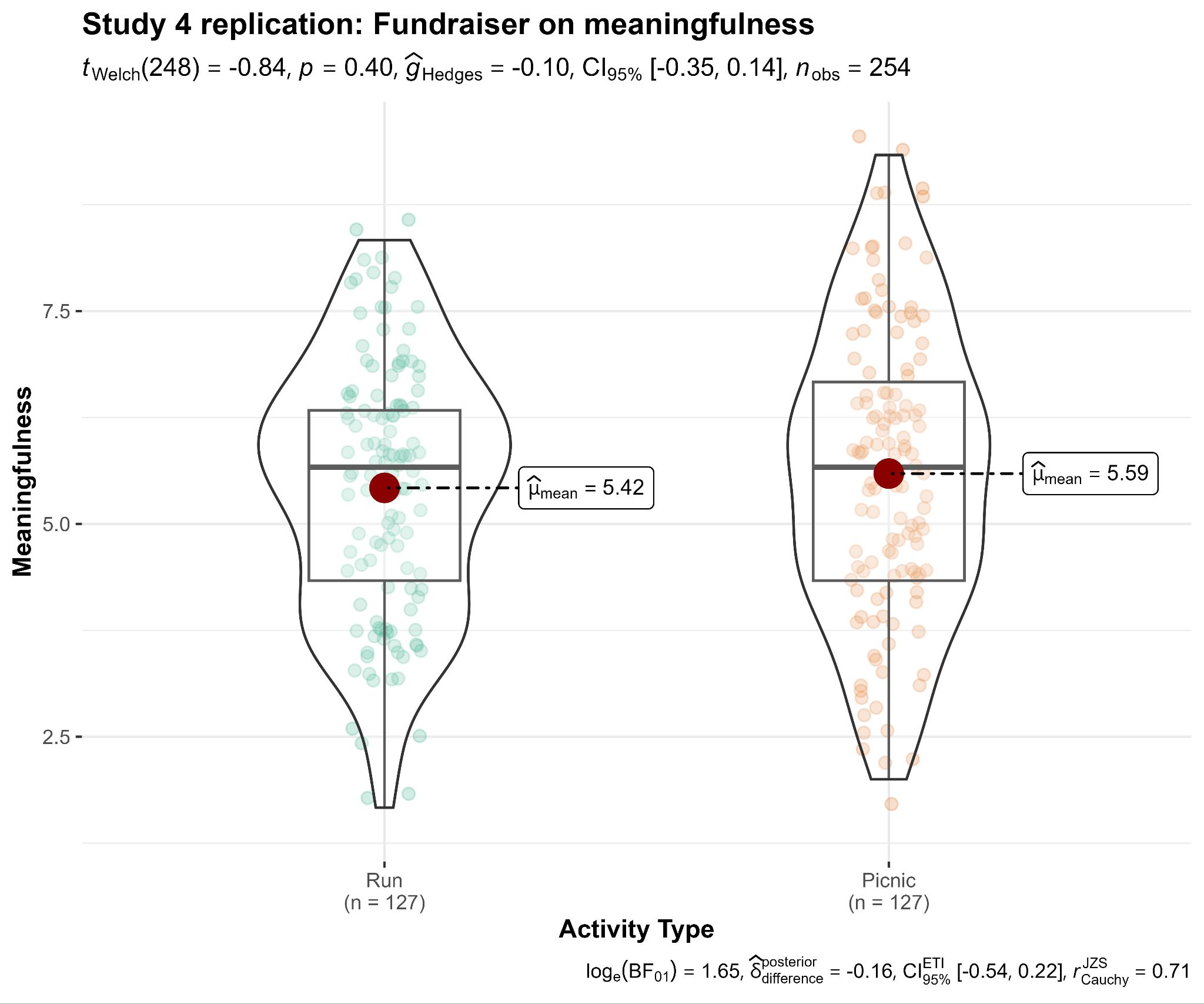


*Note*. Created using ggstatsplot (Patil, 2021)

### Meaningfulness

We conducted an independent samples t-test and found no support for difference in meaningfulness between the effortful (run) condition (*n* = 127, *M* = 5.42, *SD* = 1.49) and the easy (picnic) condition (*n* = 127, *M* = 5.59, *SD* = 1.70; *t*(252) = -0.84, *p* = 0.40; *d* = -0.10, 95% CI [-0.35, 0.14]). We provided a summary plot in Figure 3.

###### Figure 3 *Study 4 replication - meaningfulness by fundraiser type*



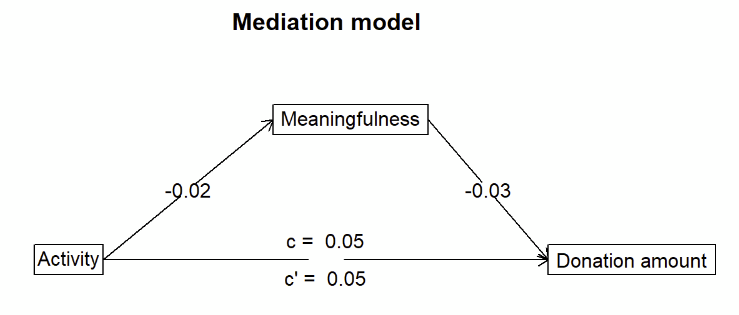
*Note*. Created using ggstatsplot (Patil, 2021)

### Meaningfulness meditation

We conducted a mediation analysis to examine if meaningfulness mediates the impact of fundraiser type (run vs. picnic) on donation amount. We found no support for the mediation. Total impact of activity on donation amount was c = 0.05 (SE = 0.04, *t* = 1.51, *p* = 0.13), with a direct effect of c' = 0.05 (SE = 0.04, *t* = 1.49, *p* = 0.14), and an indirect effect of 0 (mean bootstrapped indirect effect = 0, SE = 0, 95% CI [0, 0.01]). We provided a summary plot in Figure 4.

###### 

###### Figure 4 *Study 4 replication: Meaningfulness as mediator of impact of fundraiser type on donation amount*



## 

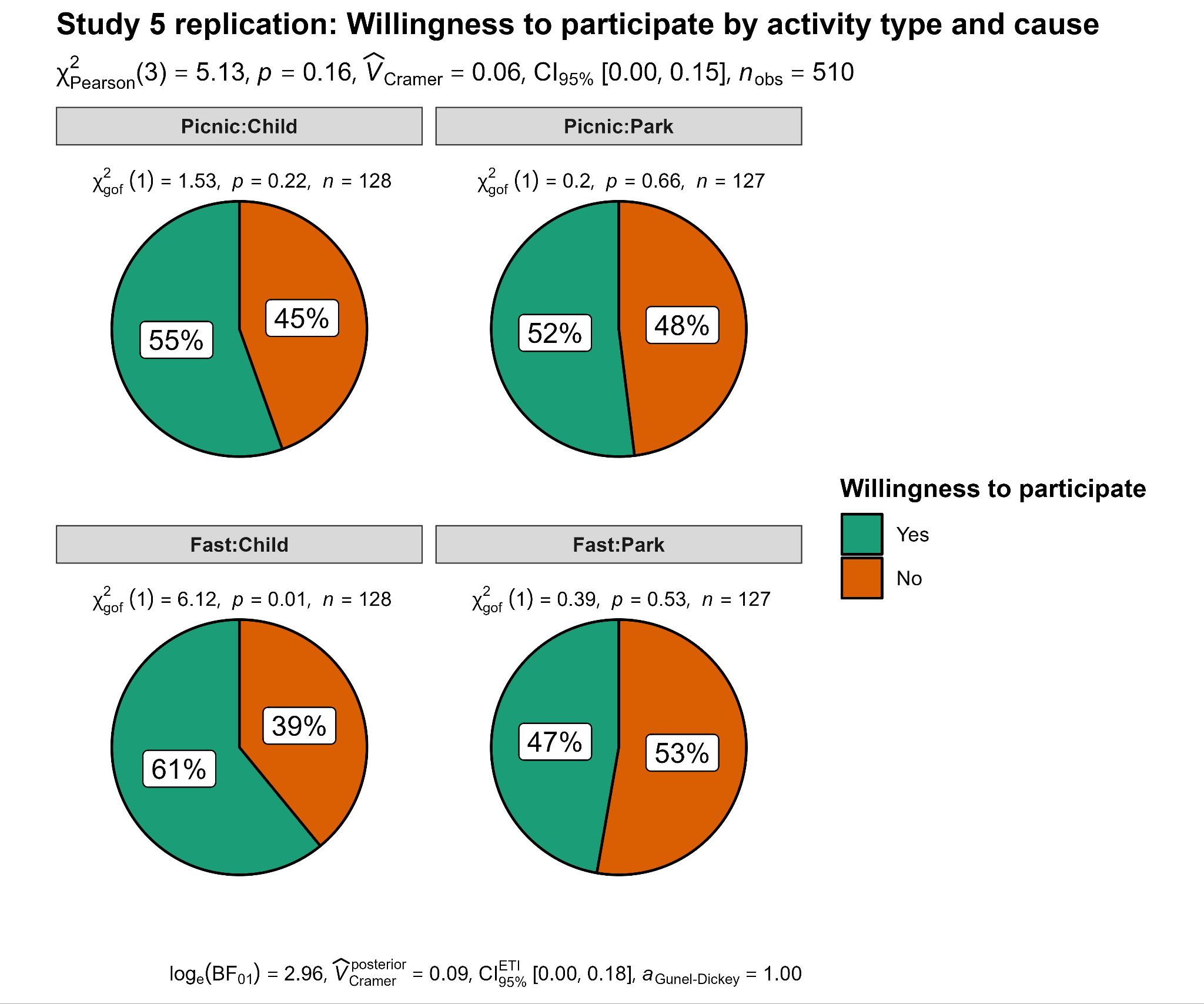
## Replication of Study 5

Mirroring the target article’s Study 5, in the unified Studies 4 and 5 design we focused on children and park causes contrasting fundraiser conditions fast and picnic.

### Willingness to participate

To mirror the target’s analysis, we conducted a chi-square test of the four conditions in the cause by activity type interaction and found no support for differences in willingness to participate between the conditions (χ²(3, *n* = 510) = 5.13, *p* = .162; Cramer's V = 0.01, 95% CI [0.00, 0.15]). We also found no support for a main effect of activity (χ²(1, *n* = 510) = .01, *p* = .929; Cramer's V = 0.00, 95% CI [0.00, 0.05]), nor for a main effect of cause (χ²(1, *n* = 510) = 3.79, *p* = .051; Cramer's V = 0.09, 95% CI [0.00, 0.17]). We provided a summary plot in Figure 5.

Figure 5   
*Study 5 replication - willingness to participate: Chisquare of four conditions of fundraiser type by cause*

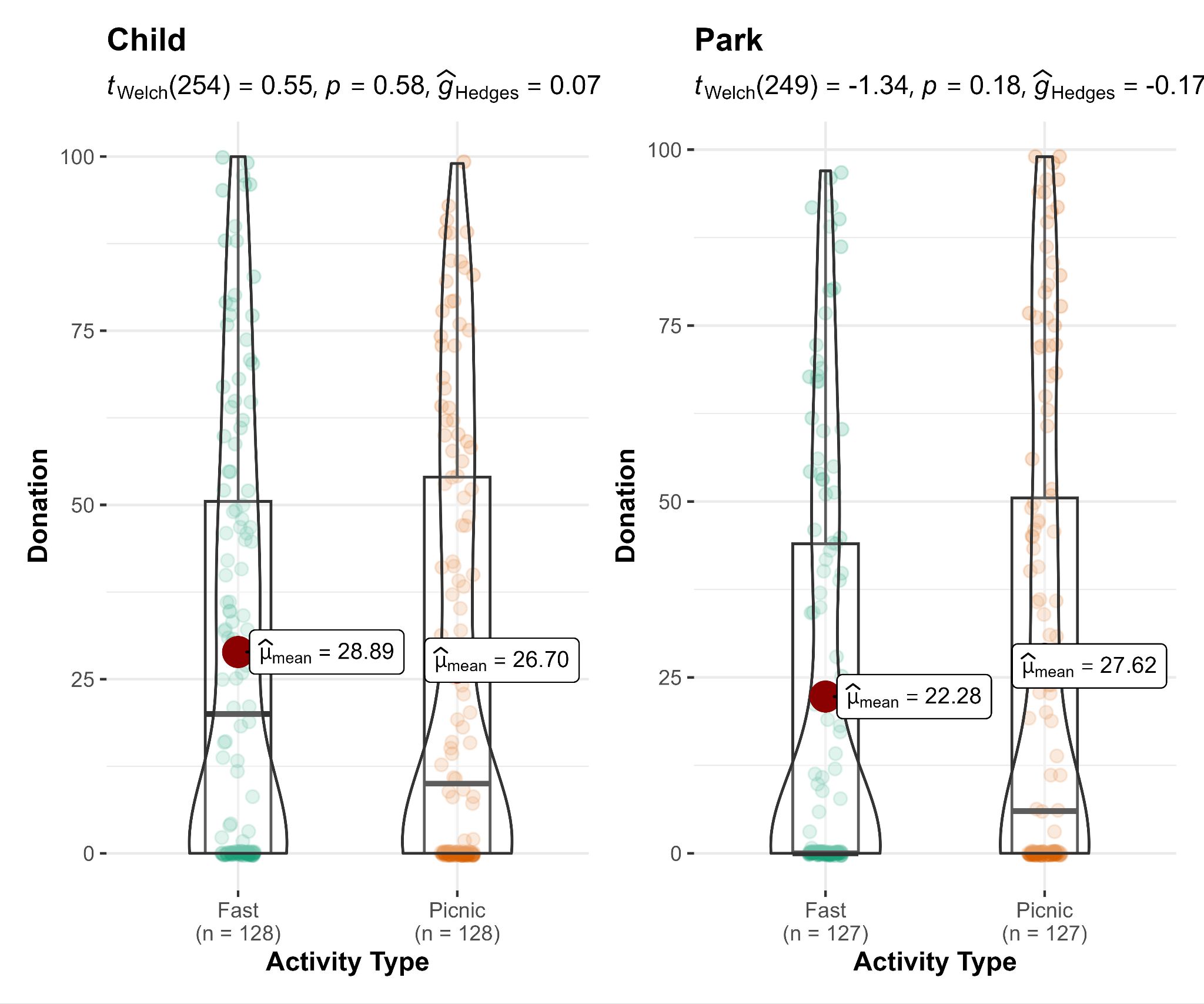


### 

### Donation amount

We conducted a two-way ANOVA to examine the interaction between fundraiser type (fast vs. picnic) and cause (children vs. park) on donation amount. We found no support for a main effect of fundraiser type (*F*(1, 506) = 0.32, *p* = .575, *η*² = .00), a main effect of cause (*F*(1, 506) = 1.03, *p* = .312, *η*² = .00) or an interaction (*F*(1, 506) = 1.80, *p* = .181, *η*² = .00). We provided a summary plot in Figure 6.

###### Figure 6 *Study 5 replication - donation amount: Interaction between fundraiser type and cause*



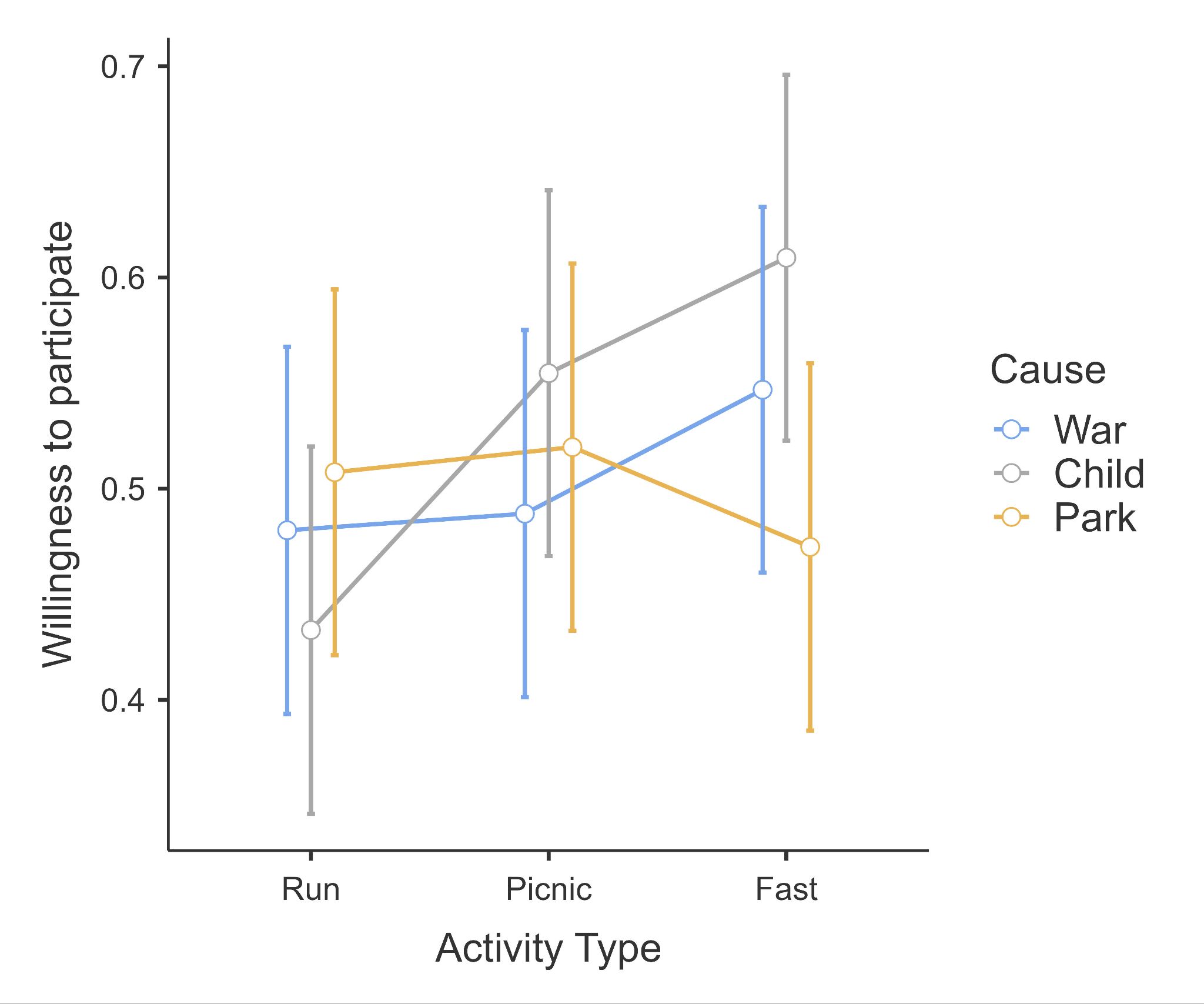
## Studies 4 and 5 unified design (extension)

### Willingness to participate

We conducted a chi-square test of the nine conditions in the cause by activity type interaction and found no support for differences in willingness to participate between the conditions (χ²(8, *N* = 1147) = 11.2, *p* = .190; Cramer's V = 0.10). We also found no support for a main effect of activity (χ²(2, *n* = 1147) = 3.83, *p* = .147; Cramer's V = 0.06), nor for a main effect of cause (χ²(2, *n* = 1147) = 0.94, *p* = .625; Cramer's V = 0.03). We provided a summary plot in Figure 7.

[Note: We will add post-hoc contrasts if any of the main effects are significant.]

###### Figure 7 *Studies 4 and 5 - willingness to participate: Interaction between fundraiser type and cause*



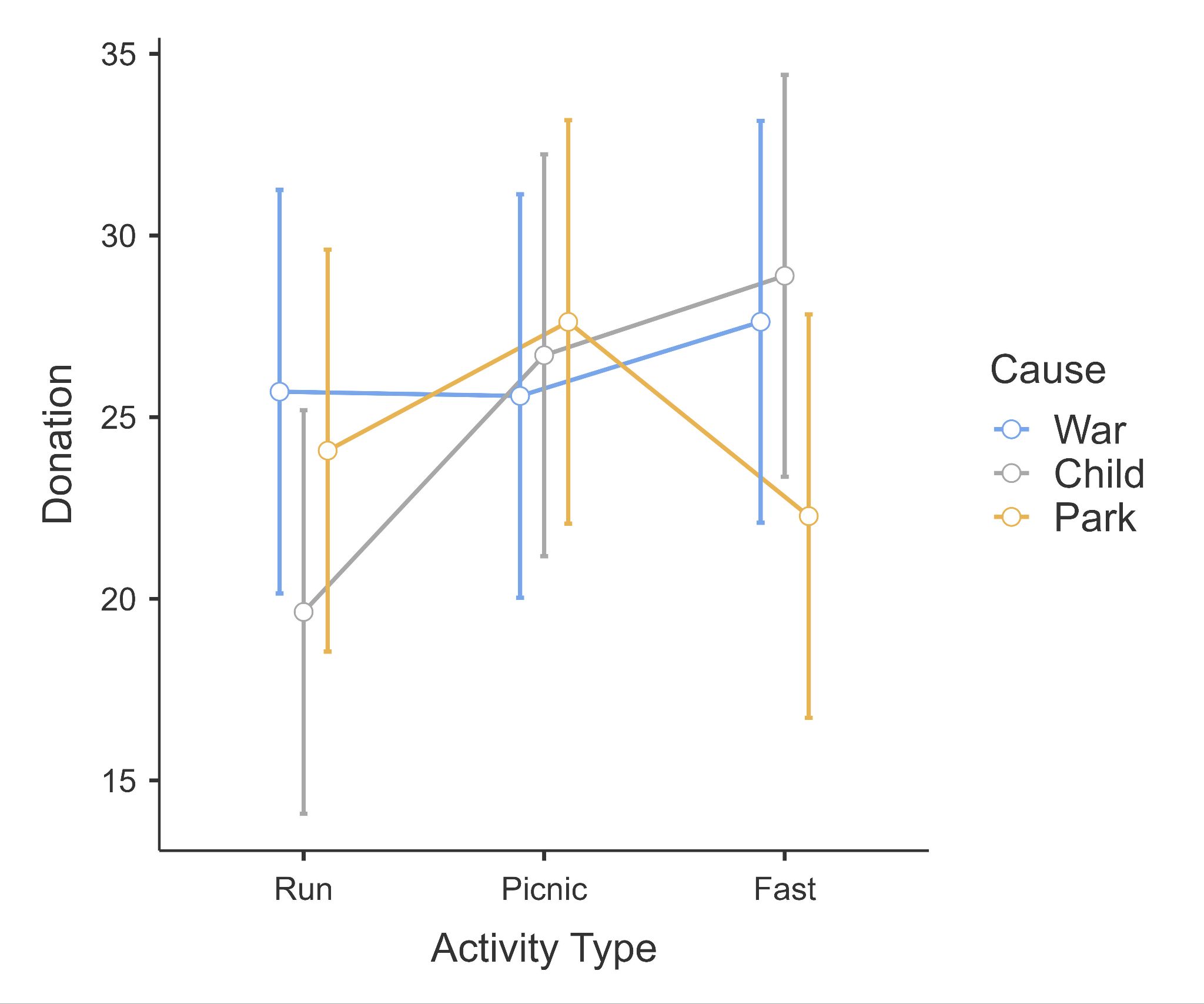
*Note*. Created with JAMOVI [Version 2.2.2] (The JAMOVI project, 2023).

### Donation Amount

We conducted a two-way ANOVA interaction of the full three fundraiser type conditions (run, picnic, fast) by three cause conditions (war, child, park) on donation amount. We found no support for a main effect of fundraiser type (*F*(2, 1138) = 1.39, *p* = .25, *η*² = 0.00), main effect of cause (*F*(2, 1138) = 0.27, *p* = .76, *η*² = 0.00), or an interaction (*F*(4, 1138) = 1.32, *p* = .26, η² = 0.01). We provided a summary plot in Figure 8.

[Note: We will add post-hoc contrasts if any of the main effects are significant.]

###### Figure 8 *Studies 4 and 5 - donation amount: Interaction between fundraiser type and cause*



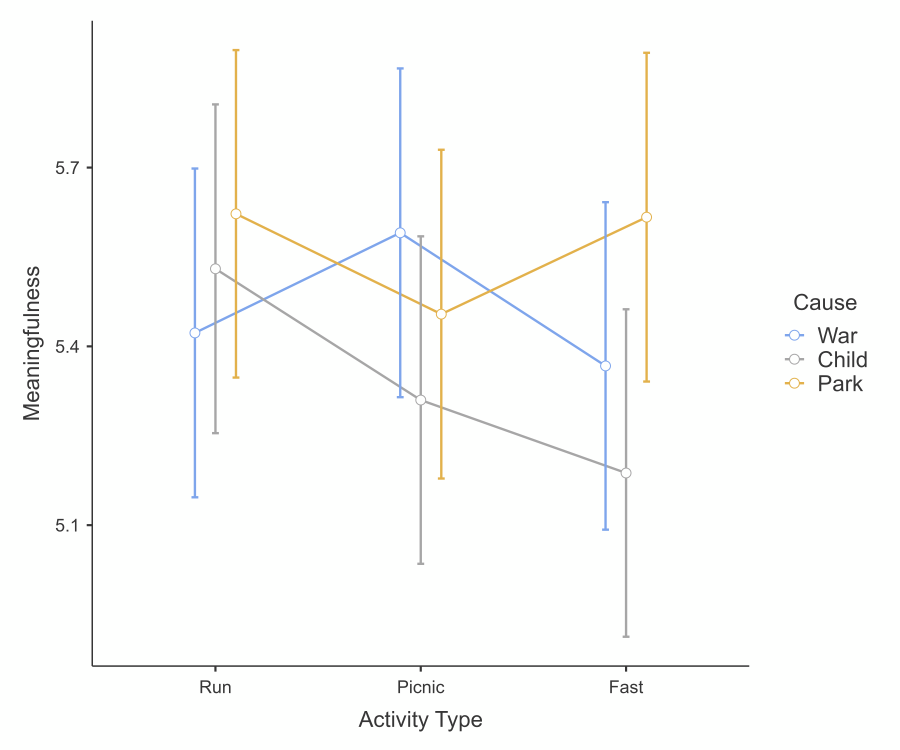
*Note*. Created with JAMOVI [Version 2.2.2] (The JAMOVI project, 2023).

### Meaningfulness

We conducted a two-way ANOVA interaction of the full three fundraiser type conditions (run, picnic, fast) by three cause conditions (war, child, park) on meaningfulness. We found no support for a main effect of fundraiser type (*F*(2, 1138) = 0.99, *p* = .41, *η*² = 0.00), main effect of cause (*F*(2, 1138) = 0.69, *p* = .50, *η*² = 0.00), or an interaction (*F*(4, 1138) = 1.88, *p* = .15, η² = 0.00). We provided a summary plot in Figure 9.

[Note: We will add post-hoc contrasts if any of the main effects are significant.]

###### Figure 9 *Studies 4 and 5 - meaningfulness: Interaction between fundraiser type and cause*

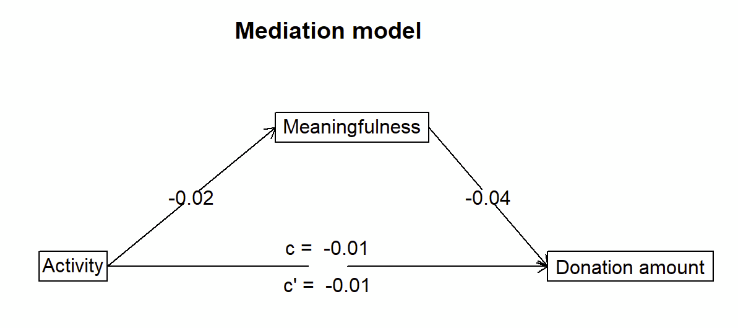


*Note*. Created with JAMOVI [Version 2.2.2] (The JAMOVI project, 2023).

### Meaningfulness meditation (fasting)

We conducted a mediation analysis to examine if meaningfulness mediates the impact of fundraiser type (fasting vs. picnic) on donation amount. We found no support for the mediation. Total impact of activity on donation amount was c = -0.01 (SE = 0.04, *t*(763) = -0.15, *p* = .88), with a direct effect of c' = -0.01 (SE = 0.04, *t*(762) = -0.18, *p* = .86), and an indirect effect of 0 (mean bootstrapped indirect effect = 0, SE = 0, 95% CI [0, 0.01]). We provided a summary plot in Figure 10.

###### Figure 10 *Studies 4 and 5: Meaningfulness mediates impact of fundraiser type on donation amount (fasting)*



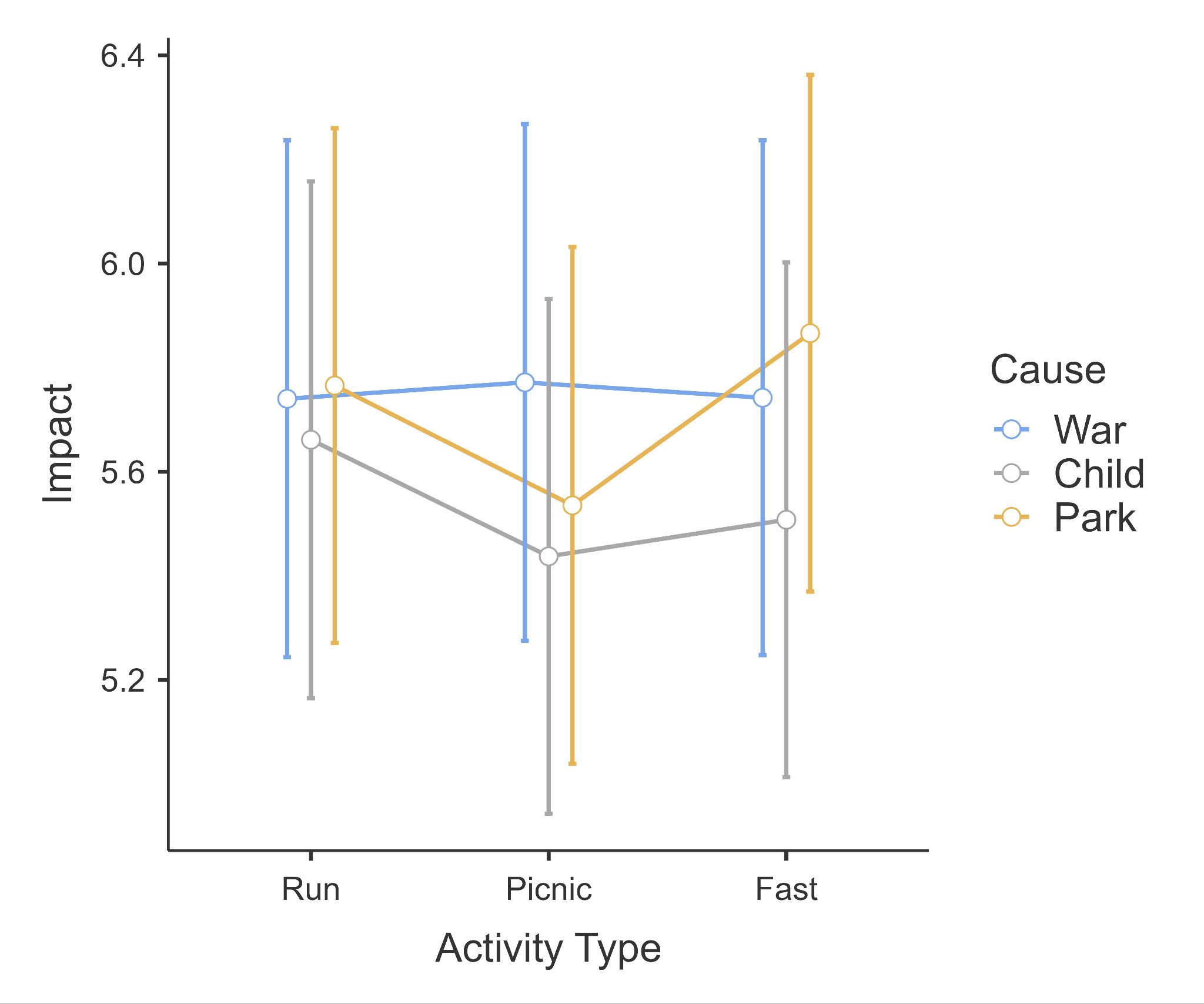
#### 

### Impact (extension)

We conducted a two-way ANOVA interaction of the full three fundraiser type conditions (run, picnic, fast) by three cause conditions (war, child, park) on impact. We found no support for a main effect of fundraiser type (*F*(2, 1138) = 0.28, *p* = .757, *η*² = 0.00), main effect of cause (*F*(2, 1138) = 0.65, *p* = .525, *η*² = 0.00), or an interaction (*F*(4, 1138) = 0.19, *p* = .943, η² = 0.00). We provided a summary plot in Figure 11.

[Note: We will add post-hoc contrasts if any of the main effects are significant.]

###### Figure 11 *Studies 4 and 5 - Impact (extension): Interaction between fundraiser type and cause*



*Note*. Created with JAMOVI [Version 2.2.2] (The JAMOVI project, 2023)

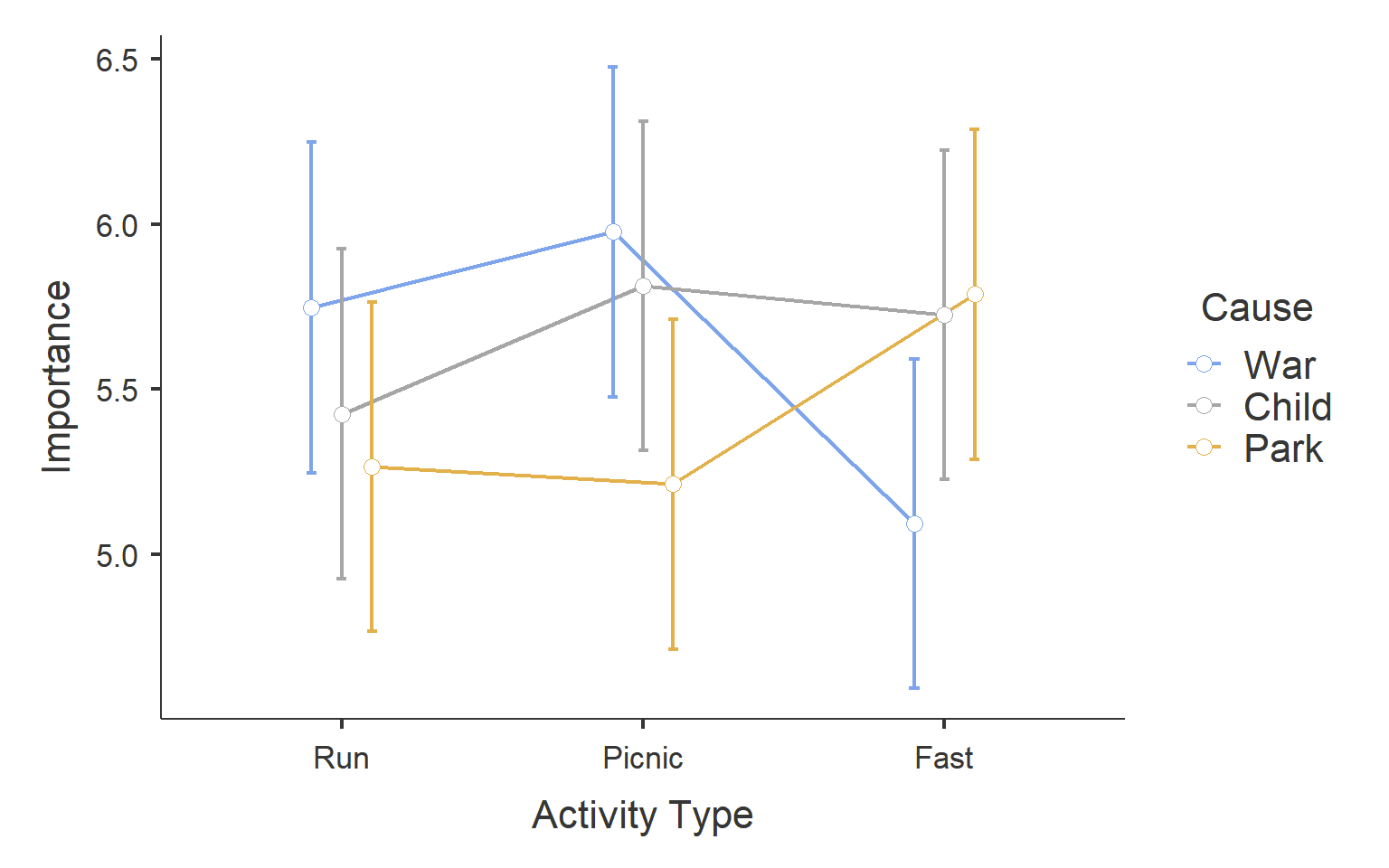
###### 

### Manipulation checks (extension)

We conducted a two-way ANOVA interaction of the full three fundraiser type conditions (run, picnic, fast) by three cause conditions (war, child, park) on the importance manipulation check. We found no support for a main effect of fundraiser type (*F*(2, 1138) = 0.43, *p* = .65, *η*² = 0.01), main effect of cause (*F*(2, 1138) = 0.70, *p* = .50, *η*² = 0.00), or an interaction (*F*(4, 1138) = 2.5, *p* = .04, *η*² = 0.01). We provided a summary plot in Figure 12.

[Note: We will add post-hoc contrasts if any of the main effects are significant.]

###### Figure 12 *Studies 4 and 5 - Importance check (extension): Interaction between fundraiser type and cause*

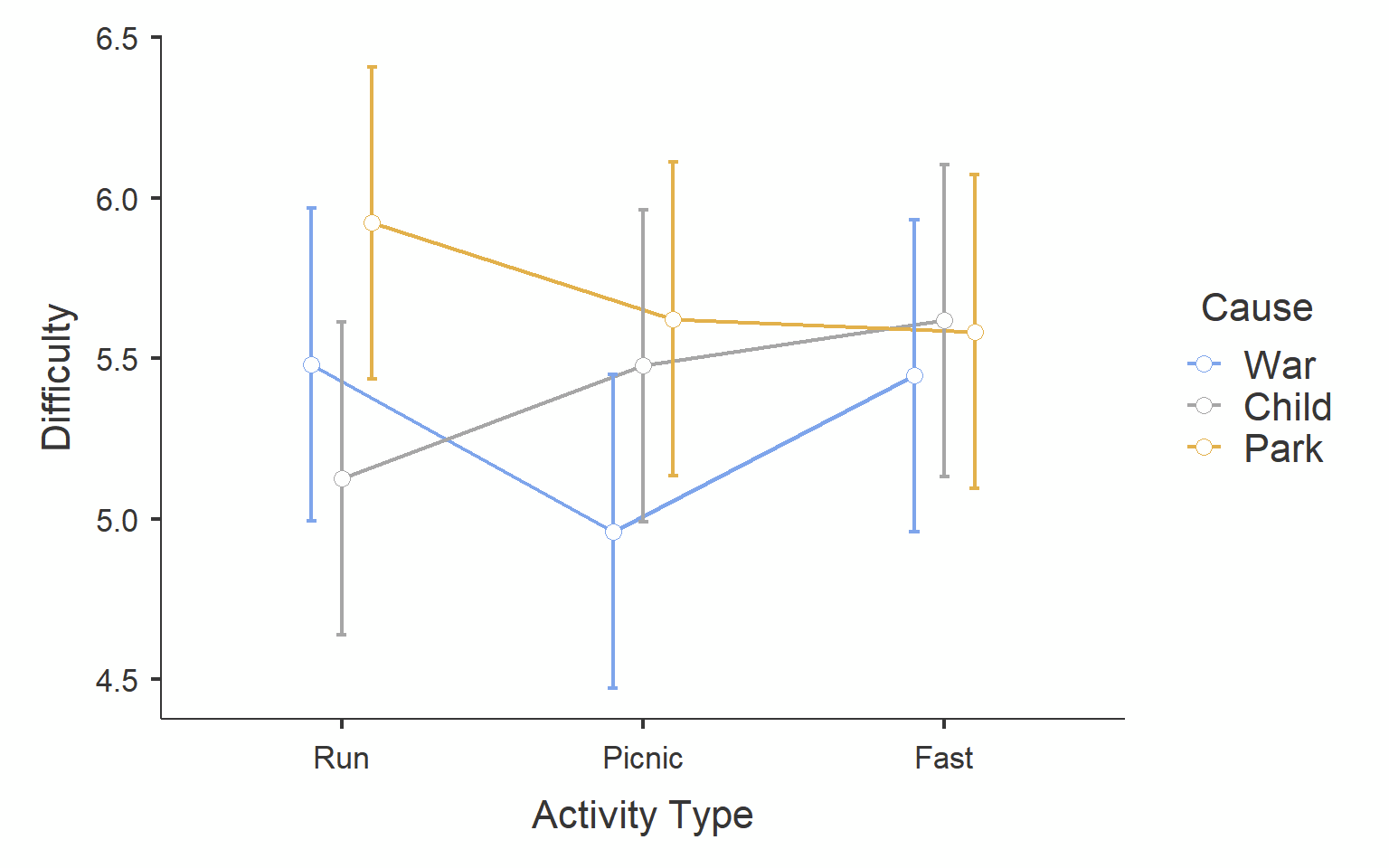


*Note*. Created with JAMOVI [Version 2.2.2] (The JAMOVI project, 2023

We conducted a two-way ANOVA interaction of the full three fundraiser type conditions (run, picnic, fast) by three cause conditions (war, child, park) on the importance manipulation check. We found no support for a main effect of fundraiser type (*F*(2, 1138) = 1.22, *p* = .65, *η*² = 0.01), main effect of cause (*F*(2, 1138) = 0.70, *p* = .50, *η*² = 0.00), or an interaction (*F*(4, 1138) = 1.22, *p* = .30, *η*² = 0.01). We provided a summary plot in Figure 13.

[Note: We will add post-hoc contrasts if any of the main effects are significant.]

###### Figure 13 *Studies 4 and 5 - Difficulty check (extension): Interaction between fundraiser type and cause*



*Note*. Created with JAMOVI [Version 2.2.2] (The JAMOVI project, 2023

### 

## Study 3

We summarized the descriptives in Table 9.

###### Table 9 *Study 3: Descriptives*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Running distance | *n* | Donation amount | *n* | Price |
| Overall | 570 | 50.83 (30.01) | 577 | 50.82 (28.91) |
| 1-mile | 29 | 39.93 (32.43) | 30 | 51.97 (29.64) |
| 2-mile | 29 | 46.62(27.03) | 30 | 59.00 (27.04) |
| 3-mile | 28 | 47.07 (29.81) | 30 | 50.00 (29.44) |
| 4-mile | 28 | 48.36 (31.60) | 28 | 52.00 (29.21) |
| 5-mile | 29 | 49.90 (33.11) | 29 | 54.62 (27.15) |
| 6-mile | 29 | 59.76 (28.12) | 29 | 48.86 (28.12) |
| 7-mile | 29 | 47.21 (31.82) | 29 | 53.97 (26.39) |
| 8-mile | 29 | 56.38 (36.69) | 29 | 53.66 (32.65) |
| 9-mile | 29 | 50.72 (31.90) | 29 | 45.83 (28.67) |
| 10-mile | 28 | 52.04 (27.43) | 28 | 45.57 (33.48) |
| 11-mile | 29 | 50.45 (30.55) | 29 | 53.97 (26.03) |
| 12-mile | 28 | 58.54 (25.05) | 29 | 54.59 (30.44) |
| 13-mile | 28 | 51.32 (34.66) | 28 | 53.07 (29.50) |
| 14-mile | 28 | 48.25 (28.03) | 28 | 40.86 (25.58) |
| 15-mile | 29 | 54.41 (27.41) | 28 | 53.07 (29.50) |
| 16-mile | 28 | 52.64 (31.17) | 30 | 48.53 (32.52) |
| 17-mile | 28 | 48.57 (28.06) | 28 | 46.25 (29.61) |
| 18-mile | 29 | 53.90 (28.71) | 29 | 50.17 (28.09) |
| 19-mile | 28 | 46.82 (27.07) | 29 | 47.28 (28.14) |
| 20-mile | 28 | 53.68 (30.03) | 28 | 52.64 (30.39) |

*Note:* Format = mean (standard deviations). *n* = indicates sample size per condition.

We conducted correlation analyses and found no support for a correlation between charity run distance and mean donation amount on an item-level (*r*(18) = .34, *p* = .16, 95% CI [-0.12, 0.68]), nor on a participant-level (*r*(568) = .05, *p* = .20, 95% CI [-0.03, .13]), also when controlling for age and gender(*r*(568) = .05, *p* = .20). We provided summary plots in Figure 14.

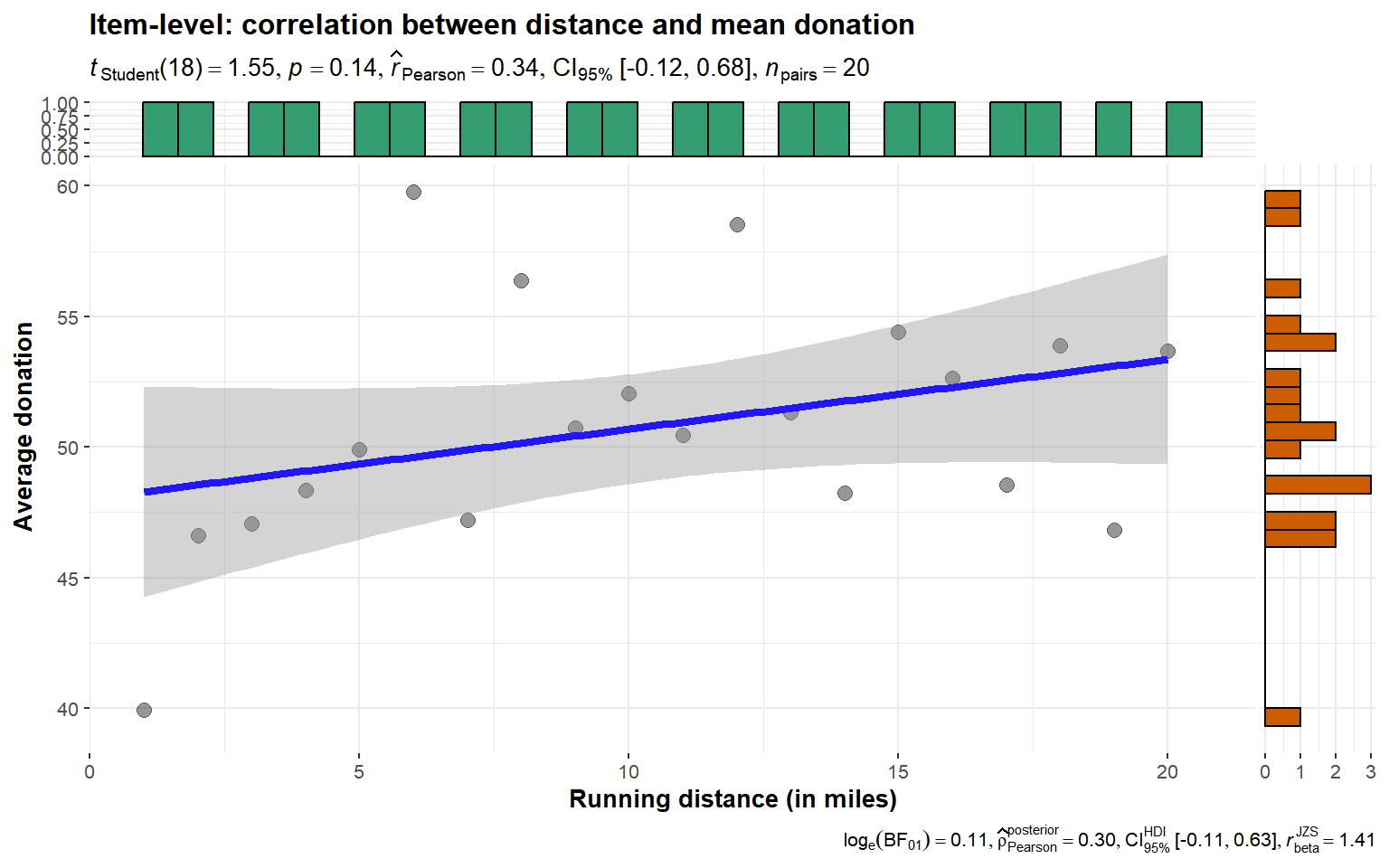
We conducted correlation analyses and found no support for a correlation between charity run distance and price required to run on an item-level (*r*(18) = -0.37, *p* = .327, 95% CI [-0.70, 0.09]), nor on a participant-level (*r*(577) = -0.05, *p* = .205, 95% CI [-0.03, .13]), also when controlling for age and gender(*r(577) = -0.05, p = .205*). We provided summary plots in Figure 15.

Comparing the correlations between the donation condition and the price condition revealed no support for differences (*z* = 1.77, *p* = .076), suggesting that the nature of the relationship between distance and the outcome (donation amount or price required) does not significantly differ across these two conditions.We provided summary plots in Figure 15.

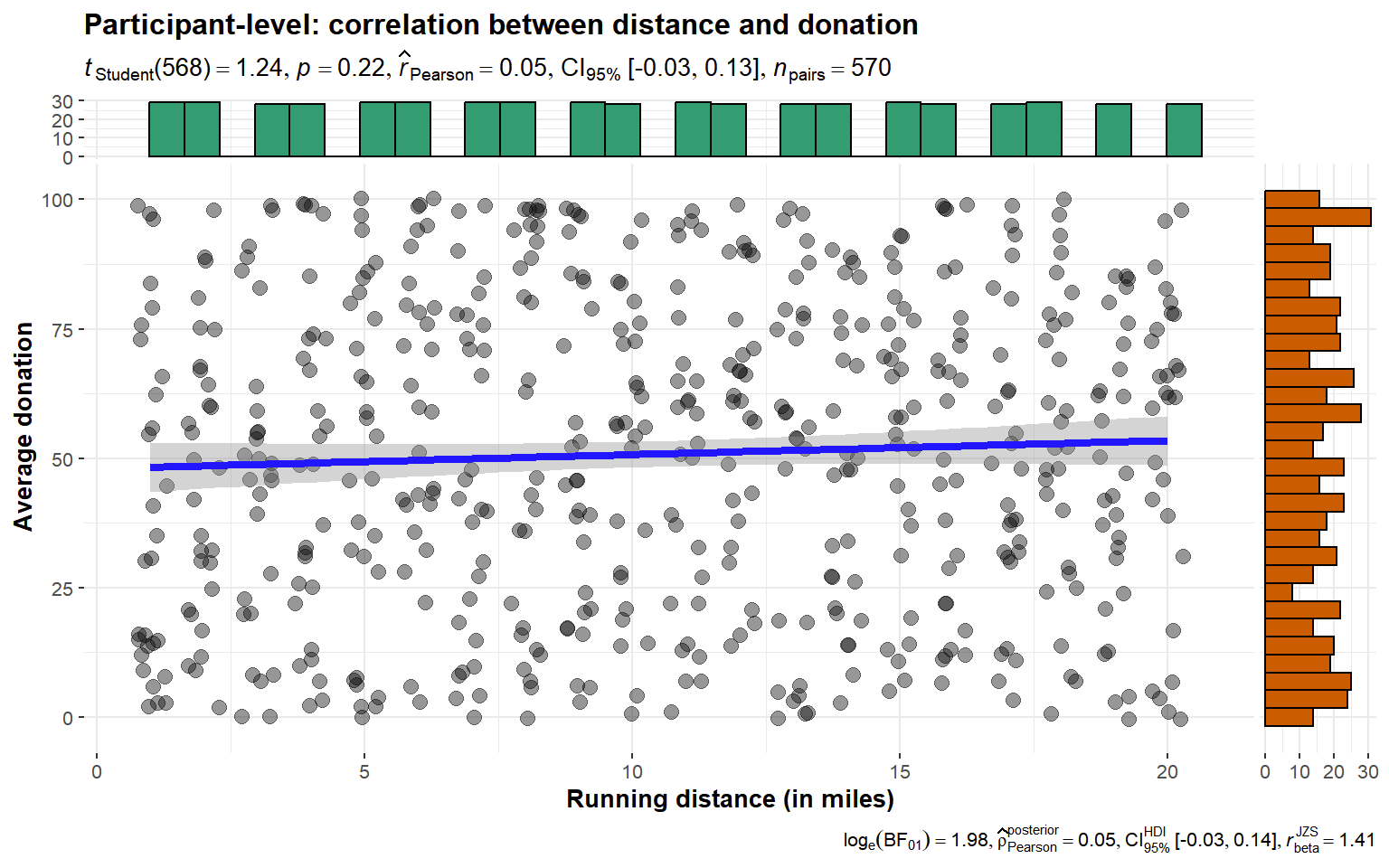
###### 

###### Figure 14 *Study 3 - donation amount: Association between charity run distance and*

Item level:



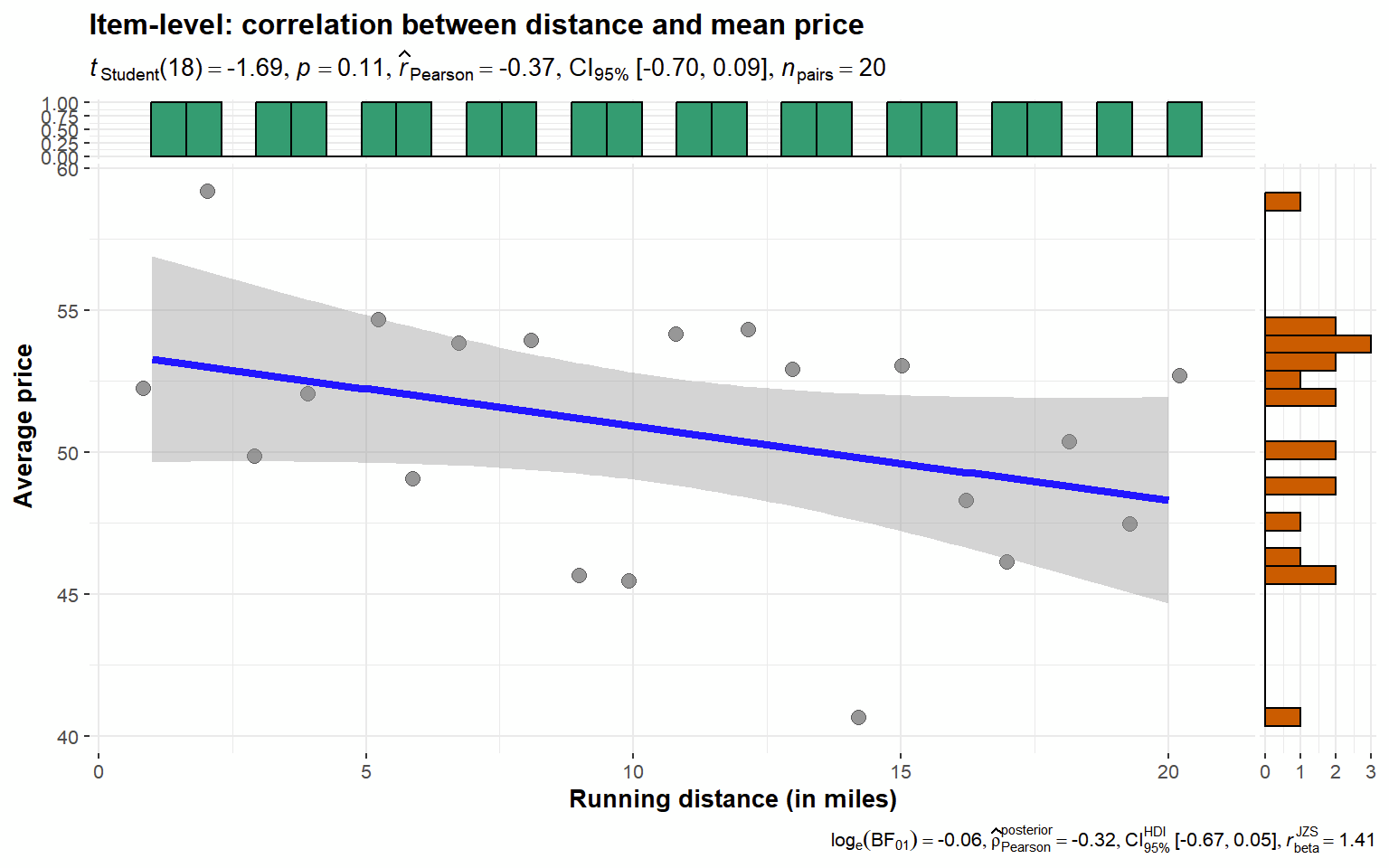
Participant level:

  
*Note*. Created using ggstatsplot (Patil, 2021)

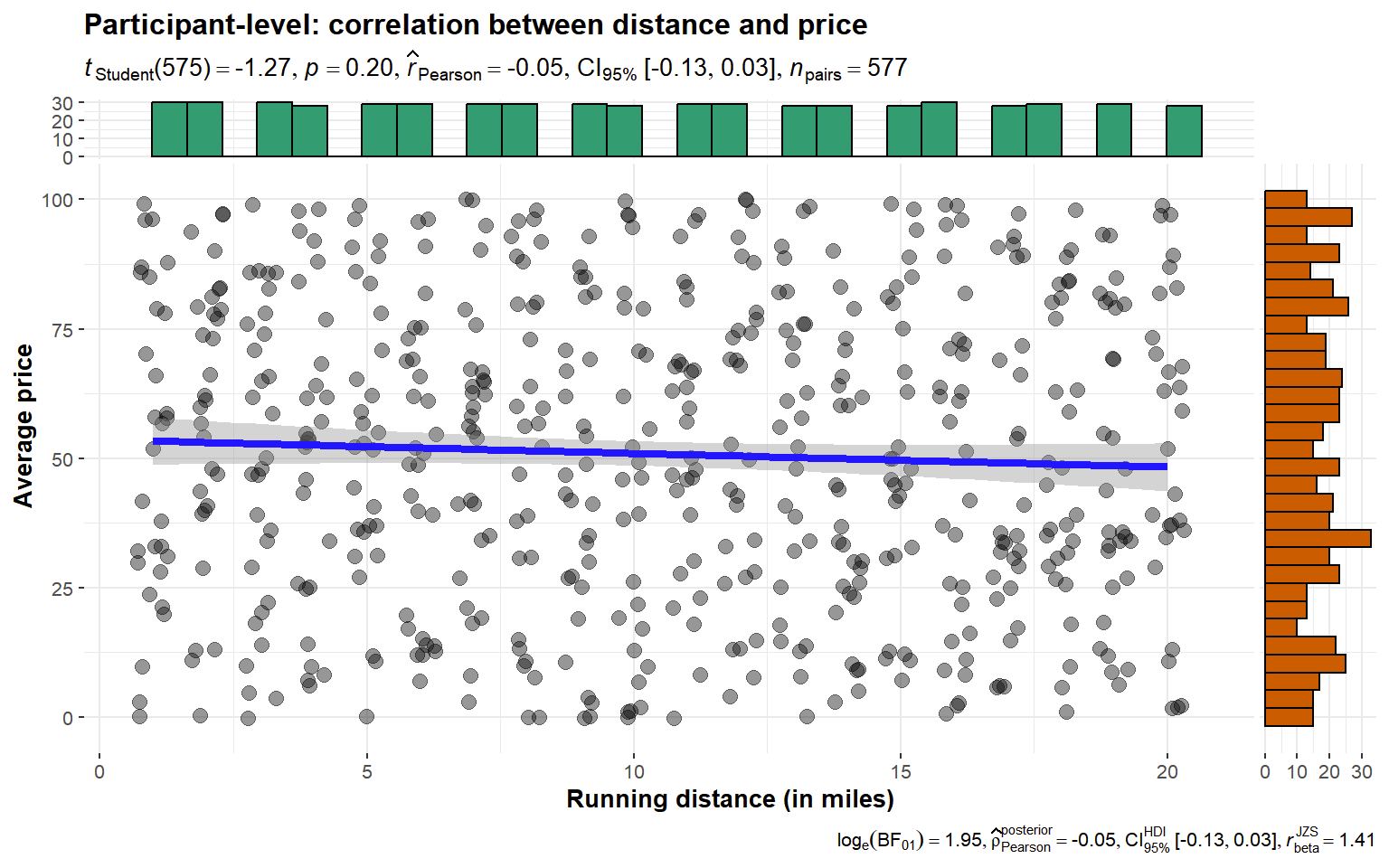
### 

###### Figure 15 *Study 3 - price: Association between run distance and*

Item level:



Participant level:



*Note*. Created using ggstatsplot (Patil, 2021)

# 

# Discussion

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

[Planned discussion for Stage 2: The target article scenarios involved matching schemes, which seem like an important factor not clearly discussed in the target’s introduction of the theory, and it is unclear whether these findings will generalize to similar fundraisers and causes with no such matching scheme. We will discuss matching schemes in relation to the theory and methodology in the target article.]

# Conclusion

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

# References

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3), 183-200. <https://doi.org/10.1037/h0024835>.

Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological science, 29*(5), 834-844. <https://doi.org/10.1177/0956797617747648>

Caviola, L., Schubert, S., & Greene, J. D. (2021). The psychology of (in) effective altruism. *Trends in Cognitive Sciences, 25*(7), 596-607. <https://doi.org/10.1016/j.tics.2021.03.015>

Caviola, L., & Greene, J. D. (2023). Boosting the impact of charitable giving with donation bundling and micromatching. *Science Advances*, 9(3). <https://doi.org/10.1126/sciadv.ade7987>

Caviola, L., Schubert, S., & Nemirow, J. (2020). The many obstacles to effective giving. *Judgment and Decision Making, 15*(2), 159-172. https://doi.org/10.1017/S1930297500007312.

Chan, M., & Feldman, G. (2024). Factors impacting effective altruism: Revisiting heuristics and biases in charity in a replication and extensions Registered Report of Baron and Szymanska (2011). Received Stage 1 in-principle acceptance from Peer Community in Registered Reports. <https://doi.org/10.17605/OSF.IO/BEP78>

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

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

Festinger, L. (1957). A theory of cognitive dissonance. Stanford University Press. <https://doi.org/10.1515/9781503620766>

Field, S. M., Hoekstra, R., Bringmann, L., & van Ravenzwaaij, D. (2019). [When and Why to Replicate: As Easy as 1, 2, 3?](https://www.collabra.org/articles/10.1525/collabra.218/). *Collabra: Psychology, 5*(1). <https://doi.org/10.1525/collabra.218>

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, 15(5), 265-268. <https://doi.org/10.1111/j.1467-8721.2006.00449.x>.

LeBel, E. P., McCarthy, R.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>

Trope, Y., & Fishbach, A. (2000). Counteractive self-control in overcoming temptation. *Journal of Personality and Social Psychology*, *79*(4), 493-506. <https://doi.org/10.1037/0022-3514.79.4.493>

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in cognitive sciences, 22*(4), 337-349. <https://doi.org/10.1016/j.tics.2018.01.007>

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., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), Heuristics and biases: *The psychology of intuitive judgment* (pp. 49-81). <https://doi.org/10.1017/CBO9780511808098.004>

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

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

MacAskill, W. (2015). *Doing good better: Effective altruism and a radical new way to make a difference*. Guardian Faber Publishing.

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>

Olivola, C. Y., & Shafir, E. (2013). The Martyrdom Effect: When Pain and Effort Increase Prosocial Contributions. *Journal of Behavioral Decision Making*, 26, 91–105. <https://doi.org/10.1002/bdm.767>

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

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).

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

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

Xygalatas, D., Mitkidis, P., Fischer, R., Reddish, P., Skewes, J., Geertz, A. W., Roepstorff, A., & Bulbulia, J. (2013). Extreme rituals promote prosociality. *Psychological Science*, *24*(8), 1602–1605. https://doi.org/10.1177/0956797612472910

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

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

1. We note that in the target article this was framed differently such that the association with donation was framed as support for no effect (null hypothesis). We reframed it to instead be about the comparison between the effect for donation and the effect for price. [↑](#footnote-ref-2)