**Revisiting the signal value of emotion in altruistic behavior:  
Replication and extension Registered Report of  
Barasch et al. (2014) Studies 3 and 6  
[Stage 1]**

Tse Lyn (Rachael) Woo   
ORCID: 0000-0002-4369-4423  
OSF: <https://osf.io/74vza/>   
Department of Psychology, The University of Hong Kong  
[rachwtl@connect.hku.hk](mailto:rachwtl@connect.hku.hk) / [rachaelwuni09@gmail.com](mailto:rachaelwuni09@gmail.com)

^Gilad Feldman  
ORCID: 0000-0003-2812-6599  
OSF: <https://osf.io/m4kfe/>   
Department of Psychology, The University of Hong Kong

[gfeldman@hku.hk](mailto:gfeldman@hku.hk) /[giladfel@gmail.com](mailto:giladfel@gmail.com)

^Corresponding author

## 

## Author bios:

Tse Lyn (Rachael) Woo was a student at the University of Hong Kong during the academic year of 2023/2024.

Dr. Gilad Feldman is an assistant professor in the Department of Psychology at the University of Hong Kong.

## 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:

Tse Lyn (Rachael) Woo 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); ORCID: 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))

## Acknowledgements (with ORCID)

We thank Prof. Barasch for providing us with the materials for Study 3.

**Important links**

Citation of the target research article:

Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of personality and social psychology*, *107*(3), 393. https://doi.org/10.1037/a0037207.

## Contributor Roles Taxonomy

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

# PCIRR-Study Design Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Question | Hypothesis | Sampling plan | Analysis plan | Rationale for tests | Interpretation given outcome | Theory/findings affected by the outcomes |
| Does emotion-driven prosocial behavior signal moral character? | 1a: Donors who exhibit high distress regarding the suffering of others, are evaluated as more moral and authentically concerned for others, compared to donors who exhibit low distress. (Replication) | We aimed to recruit 1164 US American participants online through Prolific  We aimed to recruit 1164 US American participants online through Prolific | 2 (Study 3 - Distress: High vs low; Study 6 - Expected benefit: Yes vs no ) × 3 (Study 3 - Action: Donate vs cash vs no information; Study 6: Benefit type: Reputational, material, emotional) ANOVAs with post-hoc analysis (Tukey’s HSD) for significant main effects | We followed the “small-telescope” approach (Simonsohn, 2015) and based our estimates on effects in the target article, adjusted to account for much weaker effects. Our power analysis was based on the smallest effect size observed from the omnibus two-way ANOVA results found in the target article  The effect size chosen was the smallest effect size η2p = 0.03.  Sensitivity analyses showed we should be able to detect effects of f = 0.11 (η2p = 0.01). | We examine the replicability of Barasch et al. (2014) based on the replication evaluation criteria by LeBel et al. (2019). | The study does not follow a formal theory.  However, if the replication is unsuccessful, it suggests that emotions may not serve as a positive signal of moral character in prosocial behavior |
| 2: The effect observed in the hypothesis above is stronger when a person donated compared to a person who did not donate or whose donation behavior is unspecified. (Replication) |
| 1b: Donors who exhibit high distress regarding the suffering of others, are evaluated as more authentically prosocially motivated, compared to donors who exhibit low distress. (Extension) |
| 3: A person who donated is perceived as more other-focused than a person who did not donate or whose donation behavior is unspecified.(Extension) |
| Does expecting to receive personal benefits for donations impact perceptions of donors? | 4a/b: Donors who do not expect to receive [material / reputational] benefits for their donations are perceived as more moral than donors who expect to receive [material / reputational] benefits. (Replication) |
| 4c: Donors who expect to receive emotional rewards for their donations are perceived as more moral than donors who do not expect to receive emotional benefits. (Replication) |
| 5a/b: Donors who do not expect to receive [material / reputational] benefits for their donations are perceived as more other-focused than donors who expect to receive [material / reputational] benefits. (Extension) |
| 5c: Donors who expect to receive emotional rewards for their donations are perceived as more other-focused than donors who do not expect to receive emotional benefits. (Extension) |

# 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.]

Barasch et al. (2014) demonstrated that for prosocial behavior a donating actor’s emotions can serve as a positive signal of moral character, as they are perceived as an honest and direct signal of genuine concern towards others. In a Registered Report experiment with an online Prolific sample (*N* = 1000), we conducted a replication and extension of Studies 3 and 6 from Barasch et al. (2014). [The following is a demo placeholder and will be updated following data collection.] . We [found/failed to find] support for … (XX.XX [XX, XX]). Extending the replication, we [found/failed to find] support for … (XX.XX [XX, XX]). Overall, we conclude… Materials, data, and code are available on: <https://osf.io/rwdn6/>

*Keywords:* emotion signaling, bias, judgment and decision making, registered replication, moral psychology, social perception, prosocial behavior

# 

# Revisiting the signal value of emotion in altruistic behavior: Replication and extension Registered Report of Barasch et al. (2014) Studies 3 and 6 [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

Emotions can motivate prosocial action and prosocial behavior can have an impact on one’s emotions (Batson, 2010; Cialdini et al., 1987). However, receiving benefits from a prosocial action - whether intended or not - may ‘taint’ the authenticity of the prosocial action, and whether the prosocial action was indeed to be altruistic or only meant to serve selfish goals (Batson, 2010; Cialdini et al., 1987). Barasch et al. (2014) demonstrated that lay observers consider prosocial actors more favorably when their behavior was motivated by emotions, or when benefits were strictly emotional (and not other types of benefits, such as monetary or reputational gains).

We conducted a close replication and extension Registered Report of Barasch et al. (2014) Studies 3 and 6 with the following main goals: We aimed to conduct an independent close replication of the role of emotions in influencing lay-people’s judgements of a prosocial actor’s moral character. We also aimed to extend the studies design by combining additional measures of authenticity, rewards motivation, perceived self-focus, and perceived other-focus, used in the other studies, into a single unified comprehensive design.

We begin by introducing the literature on the role of emotions in altruism and the signaling value of emotions. We then discuss our motivations for the current replication and review Barasch et al. (2014) as our chosen article for replication. We then outline our chosen studies for replication from the target article, the target’s experimental design, and our adaptations and extensions.

## On emotions in altruism

Altruism and prosocial behavior is a topic of much debate (Batson, 1987; Cialdini et al., 1973; Hobbes, 1651; Hoffman, 1981). Some have argued that if a person benefits from a prosocial action, their action might not be considered altruistic (Cialdini, 1973; 1991), whereas others argued against this interpretation, suggesting instead that it is possible for prosocial actors to have pure intentions and unintentionally benefit from their good deeds (Batson et al., 1991). The debate seems to center around the issue of intentionality and motivation behind an action, regardless of the consequences (Carlson & Zaki, 2018; Pfattheicher et al., 2022). Using this perspective, the intent of improving another’s welfare is sufficient for one’s actions to be considered prosocial (Pfattheicher et al., 2022). The crux of the matter would therefore be whether the actions were self-oriented (egotistic) or other-oriented (altruistic) (Batson & Shaw, 1991; Batson, 2011; Cialdini, 1991; Kant, 1785).

Altruistic behavior and prosocial actions are associated with emotional benefits, and are therefore considered by some to be self-interested (Cialdini, 1991; Hobbes, 1951). For example, the “Negative State Relief Model” suggests that people’s motivations in helping others are to relieve their own negative emotions evoked when witnessing another’s suffering.

Emotions may also trigger prosocial behavior, such as through empathy. The Empathy-Altruism Hypothesis (Batson et al., 1991) posits that people are motivated to act in a way that benefits others through empathic concern, or other-oriented emotions which are congruent to the emotional state felt by the other person in need (Batson, 2011). Therefore, the greater the empathic concern, the more motivated they are to help someone else (Batson, 2011).

Barasch et al. (2014) argued that motivations and intent matter, that whether people are considered prosocial actors or not is associated with perceived motivations regarding whether they sought emotional benefits, or rather that they genuinely empathized with the targets and wanted to help.

## On the signal value of emotions

Emotional expressions provide information, or send signals, to observers. They can serve as signals of the other person’s intent (Reed et al., 2012; Van Kleef et al., 2004). For example, people who tend to cooperate more, were more emotionally expressive than those who do not (Schug et al., 2010). People also seem to use emotional expressions to make inferences about a person’s traits or beliefs (e.g., Stearns & Parrott, 2012).

Emotions are also linked to morality. Emotions such as empathy, sympathy, compassion, guilt, and shame are considered to be moral emotions (Parrott, 2019; Stearns & Parrott, 2012). They may also serve as a signal of one’s moral character and affect how well-liked an individual is (Anderson et al., 2021; Stearns & Parrot, 2012). Therefore, Barasch et al. (2014) suggested that emotions may not only motivate prosocial behavior but that emotional benefits derived from a prosocial action may not be considered as negatively as reputational and material benefits.

## Choice of study for replication: Barasch et al. (2014)

We embarked on a replication and extension Registered Report of Barasch et al. (2014). We aimed to revisit the phenomenon to examine the reproducibility and replicability of the findings with an independent pre-registered well-powered close replication with adjustments and extensions. 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 Barasch et al. (2014) based on several factors: Its academic and practical impact and the potential for further extensions examining the inferences that people make when evaluating a prosocial actor.

The article has had an impact on scholarly research in the area of moral psychology and altruism. At the time of writing (January 2024), there were 301 Google Scholar citations of the article. (2014) has inspired many important follow-up theoretical and empirical articles, such as a literature on the psychology of (in)effective altruism (Berman et al., 2018; Caviola et al., 2020; Caviola et al., 2021), which is linked with the now influential Effective Altruism movement (MacAskill, 2015). Conventional views of altruism often run counter to effective altruism even when people are provided with the necessary information to work towards outcomes that maximize welfare (Berman et al., 2018). This may be driven by various factors which include but are not limited to emotional motivators (Caviola et al., 2021; Paxton et al., 2020). It also furthers research into emotion signaling which is related to other areas outside of prosocial behavior such as in cooperation (Jordan et al. 2016; Levine et al., 2018) and decision making (VanBergen et al., 2022).

Barasch et al. (2014)'s work has important practical implications, aiding in the understanding of prosocial altruistic behavior and how it is perceived by observers, to try and better align people’s goals and reality and address misunderstandings or misalignments (Berman & Silver, 2022; Epley et al., 2023; Klein et al., 2015). Despite its impact, to our knowledge, there are currently no published direct replications of their study.

## 

## Chosen studies for replication: Studies 3 and 6

We focused our investigation on the studies examining whether levels of distress and subsequent behavior affected how lay people perceived the target actor and whether expecting to receive a benefit and the type of benefit affected judgements of the prosocial actor: Studies 3 and 6. We chose not to replicate studies 1, 2, 4, and 5. Studies 1 and 2 broadly examined how the intensity of emotion and the type of emotions that motivates prosocial behavior affect how a prosocial agent is perceived. Studies 4 and 5 examined how prosocial agents were perceived when they received emotional benefits. Most of the studies in the target article employed similar measures and Studies 3 and 6 built upon the findings of their predecessors (Studies 1 and 2 and Studies 4 and 5 respectively), and we therefore decided to focus on Studies 3 and 6. We provided a summary of Studies 1, 2, 4 and 5’s hypotheses in the “Analysis of the Target Article” section of the supplementary document.

We summarized the setup of Studies 3 and 6, along with the corresponding hypotheses and findings from the target article in Tables 1 and 2.

## Barasch et al. (2014): Hypotheses and findings

We provided a summary of the hypotheses of the target article in Table 1. The main hypothesis of the article was that emotional benefits derived from prosocial behavior signal stronger moral character to lay observers. We listed the hypotheses of Studies 3 and 6 in Table 1 and summarized their associated findings in Table 2.

###### 

###### **Table 1** *Barasch et al. (2014) Studies 3 and 6: Summary of hypotheses*

|  |  |  |  |
| --- | --- | --- | --- |
| **S** | **Scenario** | **H#** | **Hypothesis** |
| 3 | Prosocial versus selfish behavior | 1a | (Donation condition only)  Donors who exhibit high distress regarding the suffering of others, are evaluated as more moral and authentically concerned for others, compared to donors who exhibit low distress. (Core main effect hypothesis, similar to Studies 1 and 2, see supplementary) |
|
|
| 2 | Interaction between emotions and behavior in predicting perceived moral character and authenticity. That is: Main effect in H1a is stronger when a person donated compared to a person who did not donate or whose donation behavior is unspecified. |
| Extension | 1b | (Donation condition only)  Donors who exhibit high distress regarding the suffering of others, are evaluated as more authentically prosocially motivated, compared to donors who exhibit low distress. |
| 3 | A person who donated is perceived as more other-focused than a person who did not donate or whose donation behavior is unspecified. |
| 6 | Type of rewards and anticipation of rewards | 4a | Donors who do not expect to receive material benefits for their donations are perceived as more moral than donors who expect to receive material benefits. |
| 4b | Donors who do not expect to receive reputational benefits for their donations are perceived as more moral than donors who expect to receive reputational benefits. |
| 4c | Donors who expect to receive emotional rewards for their donations are perceived as more moral than donors who do not expect to receive emotional benefits. |
| Extension | 5a | Donors who do not expect to receive material benefits for their donation are perceived as more other-focused than donors who expect to receive material benefits. |
| 5b | Donors who do not expect to receive reputational benefits for their donation are perceived as more other-focused than donors who expect to receive reputational benefits. |
| 5c | Donors who expect to receive emotional rewards for their donations are perceived as more other-focused than donors who do not expect to receive emotional benefits. |

*Note*: Some of the hypotheses were not explicitly stated and were deduced from the target article’s introduction and design.

**Table** **2***Barasch et al. (2014) Studies 3 and 6: Summary of findings*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Experiment 3 (*N* = 470)** | | | | | | |
| **DV** | **IV** | ***F* statistic** | ***p*** | ***η2p*** | **CIL** | **CIH** |
| Distress manipulation check | Distress | *F*(1, 464) = 432.68 | <.001 | 0.48 | 0.42 | 0.54 |
| Distress x action | *F*(2,464) = 3.04 | .05 | 0.01 | 0.00 | 0.04 |
| Moral character | Distress | *F*(1, 464) = 51.56 | <.001 | 0.10 | 0.05 | 0.15 |
| Distress x action | *F* (2,464) = 12.73 | <.001 | 0.05 | 0.02 | 0.09 |
| **Experiment 6 (*N* = 286)1** | | | | | | |
| **DV** | **IV** | ***F* statistic** | ***p*** | ***η2p*** | **CIL** | **CIH** |
| Moral character | Expected benefit | *F*(1, 280) = 34.68 | <.001 | 0.11 | 0.05 | 0.18 |
| Expected x benefit type | *F*(1,280)=25.43 | <.001 | 0.15 | 0.08 | 0.23 |
| Authentic prosocial motivation | Expected benefit | *F*(1, 280) = 68.59 | <.001 | 0.20 | 0.12 | 0.28 |
| Benefit type | *F*(1,280)=7.46 | .001 | 0.03 | 0.00 | 0.07 |
| Expected x benefit type | F(1,280) = 25.82 | <.001 | 0.16 | 0.08 | 0.23 |

*Note*: CIL = lower bounds for CIs. CIH = higher bounds of CIs. IV = Independent variable; DV = Dependent variable.  
Effect sizes and confidence intervals were not reported in the target and are based on our calculations reconstructed from the target article’s reported statistics.   
1 The reported sample size in Experiment 6 was 276, though degrees of freedom reported in the results did not add up to 276 (pp. 406-407).

## Extensions

We used the designs of Studies 3 and 6, the two chosen studies, as a base for adding extensions, adopting the measures used in Studies 1, 2, 4, and 5 in Barasch et al. (2014) with adjustments, aiming for a unified comprehensive examination of many factors studied in the target article within a single design.

### Study 3

We aimed to extend the replication study by adding measures of authentic prosocial motivation, perceived self and other-focus of the targets’ actions, and emotional benefits motivation . We considered authentic prosocial motivation and perceived self and other focus as core extensions due to their relationship with perceived moral character, and the ongoing debate between the empathy-altruism hypothesis and negative state relief theory. Authentic prosocial motivation was not originally included in the study because the original items listed would only be applicable to the donation condition (p. 401). However, we thought it relevant to try and determine whether or not lay observers thought there were differences in perceived authentic concern for others when donating, when not donating, and when the decision was ambiguous. In line with the findings of the other studies in the target article, we posited that targets experiencing high distress in the donate condition will be seen as more authentically motivated than targets experiencing low distress in the donate condition. Furthermore, given that targets in the control condition were seen as more moral than targets in the no-donate condition (p. 402) we also posited that targets in the control condition would be seen as more authentically motivated than targets in the cash (no-donate) condition.

We also set out to examine the extent to which lay observers thought the targets’ actions were self or other focused, to connect to the existing literature about the empathy-altruism hypothesis and the negative state relief theory. Study 2 of the target article found that participants perceived targets who experienced higher emotional arousal to be more other-focused than targets who experience less emotional arousal (p. 400), and so we hypothesized that we would see a similar result here such that targets in the high-distress donate condition would be seen as more other-focused than targets in the low-distress donate condition. Additionally, we expected that targets in the donate and control conditions would be seen as more other-focused than targets in the cash condition, and vice versa, such that targets in the cash condition would be seen as more self-focused than both the donate and control conditions.

### Study 6

We aimed to extend the replication study by adding measures of perceived self and other focus. The target study found that expecting emotional benefits increased perceptions of moral character while expecting reputational or material rewards had a negative impact on moral character. We therefore set out to determine if lay observers considered expecting an emotional benefit to be more other-focused than expecting other types of benefits. We predicted that donors who expect to receive emotional benefits are perceived as more other-focused than donors who expect to receive material and reputational benefits. Similarly, we also predicted that donors who do not expect benefits are perceived as more other-focused than donors who expect to receive a benefit, regardless of benefit type.

## Pre-registration and open-science

We provided all materials, data, and code on: <https://osf.io/rwdn6/>. This project received Peer Community in Registered Report Stage 1 in-principle acceptance ((ENTER LINK AFTER IPA); (ENTER LINK AFTER IPA)) after which we created a frozen pre-registration version of the entire Stage 1 packet (ENTER LINK AFTER IPA) and proceeded to data collection. All measures, manipulations, exclusions conducted for this investigation are reported, and data collection was completed before analyses. This Registered Report was written using the Registered Report template by Feldman (2023).

# Method

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

## Power and sensitivity analyses

We first calculated effect sizes (ES) and conducted a power analysis based on the effects 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 “effectsize” (Version 0.8.6; Ben-Shachar et al., 2020), “MOTE” (Version 1.0.2; Buchanan et al., 2017), and “pwr” (Version 1.3-0; Champely et al., 2017) packages. Power analysis was conducted using GPower (Version 3.1; Faul et al., 2007) for the factors that the authors found support for in the target article (flagged as significant results).

Rounding up to the highest minimum sample size required for both studies, we concluded that the minimum required sample size was 423 participants in total.‎ This calculation is based on the effect size of *η2p* = 0.03 with a power of 0.95 and alpha of 0.05. This number was chosen because it was the smallest effect size for a supported hypothesis reported in the omnibus ANOVA test (effect of benefit type on authentic prosocial motivation in Study 6). We provided more information regarding these calculations in the “Power analysis of the original study effect to assess the required sample for replication” subsection of the supplementary materials.

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 (*d*33%) with the general rule of thumb to multiply the estimated required sample of 423 by 2.5, even if meant for other designs. This resulted in a sample of 1058. Accounting for possible exclusions of 10% based on our previous experience with the target sample, our integrated design, and allowing for the potential of additional analyses, we aimed for a larger total sample of 1164 participants, over 1.54 times larger than the combined samples in the target article (target article’s Study 3 had 470 participants and Study 6 had 286 participants). A sensitivity analysis using Gpower (Faul et al., 2007) indicated that a sample of 1058 (after exclusions) would allow the detection of *f* = 0.11 (i.e., *η2p* = 0.01) for a 2×3 two-way ANOVA for our experimental design (95% power, alpha = 5%, one-tail). These are much smaller effects than those reported in the article

## Participants

[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 1164 and the real data.]

We recruited a total of 1000 US American participants through Prolific (*Mage* = 50.8 SD = 29.9, 237 females, 247 males, 516 other or 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”, “Place of most time spent before turning 18: United States”, “Minimum Approval Rate: 95, Maximum Approval Rate: 100”, “Minimum Submissions: 50, Maximum Submissions: 100000”.

[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.]

[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 analyzed other than to assess technical issues, survey completion duration, and needed pay adjustments, and were included in the final data analysis.]

###### Table 3 *Difference and similarities between original study and replication*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Barasch et al.  Study 3 | Barasch et al.  Study 6 | Replication  US Prolific |
| Sample size | 460 | 286 | 1000 | |
| Geographic origin | Unreported | Unreported | US American | |
| Gender | 165 females, unknown number of males and other/did not disclose | 119 females, unknown number of males and other/did not disclose | 247 males, 237 females, 516 other/did not disclose | |
| Median age (years) | Unreported | Unreported | 51.5 | |
| Average age (years) | 28.0 | 34 | 50.8 | |
| Standard deviation age (years) | Unreported | Unreported | 29.9 | |
| Age range (years) | Unreported | Unreported | 0 - 100 | |
| Medium (location) | Computer (online) | Computer (online) | Computer (online) | |
| Compensation | Unreported | Unreported | Nominal payment | |
| Year | 2014 or earlier | 2014 or earlier | 2024 | |

*Note:* In the target article, the authors noted that 35% participants were female in Study 3, and 41.3% were female in Study 6.

## Experimental design

In the target article, Studies 3 and 6 were conducted separately with independent samples. We ran both studies together in a single unified data collection. The display of scenarios and conditions was counterbalanced using the randomizer “evenly present” function in Qualtrics. Participants were either randomly assigned to complete Study 3 first and then Study 6 or vice versa. 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 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 summarized the experimental design for Study 3 in Table 4 and for Study 6 in Table 5, and our adjustments to the target article in Table 6. Both studies employed a 2×3 between-subjects design. Study 3 was a 2 (Distress: High vs low) by 3 (Action: Donate, cash, vs no information) design. Study 6 was a 2 (Expected: Yes vs no) by 3 (Benefit type: Reputational, material, vs emotional) design. The presentation of conditions was counterbalanced using the “randomiser” and “evenly present” functions in Qualtrics. A combination of new comprehension checks and existing ones taken from the article were used to ensure that participants fully understood the instructions before proceeding to the actual questions.

## Procedure and materials

[*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/b5865bb6-25bb-4ef7-8bda-3cdf155f698e/SV_b4zbijNpv158c86?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. We also reached out to the authors of the target article and are very grateful for the materials for Study 3 provided by the lead author 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).

Participants were then randomly assigned into either one of six conditions in Study 3 or Study 6 through the “evenly present” and “randomiser” functions on Qualtrics, and again into the second study (either Study 3 or Study 6 depending on which study has not been completed yet).

There were six conditions in each study: In Study 3, participants in the donate condition (*n* = 334) were told that the galvanic skin response records (either high distress (*n* = 167) or low distress (*n* = 167) as indicated by a slider) they were viewing were of subjects in a previous experiment and that they opted to donate the bonus they received from participating in said experiment to charity. Participants in the cash condition (*n* = 333) were shown similar stimuli (high distress, *n* = 166 vs low distress, *n* = 167) but were told that participants opted to take the bonus for themselves in cash. Participants in the control condition (*n* = 333) were also shown similar stimuli (high distress, *n* = 166 vs low distress, *n* = 167) but they were not given information about what the subject did with the money (i.e., unclear whether they donated or took the cash).

In Study 6, all participants viewed a scenario in which a person named Jeff read a newspaper article and then decided to donate to a local soup kitchen. Participants in the reputational benefits condition (*n* = 333) were told that Jeff either expected (*n* = 166) or did not expect (*n* = 167) to receive public recognition for his contribution. Participants in the material benefits condition (*n* = 333) were told that Jeff either expected (*n* = 167) or did not expect (*n* = 166) to receive a tax break after donating. Participants in the emotional benefits condition (*n* = 334) were told that Jeff either expected (*n* = 167) or did not expect (*n* = 167) to be happier after making the donation.

We note that when reconstructing the materials we noticed that the conditions in Study 6, at least as described in the target article, were not entirely equivalent, and seemed to conflate expectations and outcome. For example, expectations for reputation seemed to conflate whether the donation was private or public. We categorize this as a possible weakness in the experimental design and decided to deviate and make an adjustment to the target’s stimuli to focus solely on manipulation of expectations.

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

###### Table 4 *Study 3: Replication and extension experimental design*

|  |  |  |  |
| --- | --- | --- | --- |
| **IV1: Action**  **IV2: Distress**  (between subject) | **IV1: Donation**  Donated the bonus $3 to the African Children's Fund | **IV1: Kept cash**  Donor opted to receive the $3 in cash | **IV1: Control condition**  No information on what donor did |
| **IV2: High distress**  Indicated by GSR level | **Dependent variables**  Moral character [Replication]  Please rate the extent to which you agree that the following words fit the person  [12-item scale w/ 6 positive (moral, altruistic, sincere, pure, good, nice) and 6 reverse-coded traits (immoral, selfish, insincere, impure, bad, mean)]  (1 = *Not at all* to 7 = *Extremely)*  Inferred emotion type [Replication]  To what extent do you think the person felt the following [options: sympathy, distress, sadness, happiness, guilt, anger, pity, discomfort, compassion]  (1 = *Not at all* to 7 *Very much*)  Emotion-level manipulation [Replication]  To what extent do you think the person was motivated by feelings?  (1 = *Strongly agree* to 7 *Strongly disagree*)  Likelihood of donating (no info ONLY) [Replication]  “How likely is it that this person donated money to the African Children’s Fund?”  (1 = *Not at all likely* to 7 *Extremely likely*)  Authentic prosocial motivation [Extension]  “How genuine do you find the person’s intentions?”  (1 = *Not at all* to 7 *Extremely*)  “To what extent is the person’s behavior reflective of their intentions?”  (1 = *Not at all* *reflective* to 7 *Extremely reflective*)  “To what extent do you agree with the following statements:”  - “The person has a genuine passion for the Africa's Children Fund”  - “The person sincerely cares about children in Africa”  - “The person wants to help children in Africa”  (1 = *Strongly agree* to 7 *Strongly disagree*)  Self-focus [Extension]  “To what extent was the person thinking about him/herself?”  (1 = *Not at all* to 7 *Very much*)  Other-focused [Extension]  “To what extent was the person thinking about others”  (1 = *Not at all* to 7 *Very much*)  Emotional benefits motivation [Extension]  “To what extent do you agree with the following statement”   * “The donor donated to the African Children’s Fund to make himself feel better” (Donate condition) * “The donor opted for taking the cash to make himself feel better” (Cash condition; reverse-coded) * “The person wanted to make him/herself feel better” (Control condition; reverse-coded)   (1 = *Strongly agree* to 7 *Strongly disagree*)  **Comprehension checks**  With condition matched validation checks.  **“**What were the person’s galvanic skin responses?”  (*Low*; *neutral*; *high*)  What does the galvanic skin response measure?”  (*Sympathy*; *distress*; *it doesn’t say*)  “How much did the person donate?”  (*$0 - did not donate*; *Donated $3*; *it doesn’t say*) | | |
|
|
| **IV2: Low distress**  Indicated by GSR level |
|
|

###### 

###### Table 5 *Study 6: Replication and extension experimental design*

|  |  |  |  |
| --- | --- | --- | --- |
| **IV1: Benefit type**  **IV2: Expected benefits**  (between subject) | **IV1: Emotional benefit**  Benefit = happiness | **IV1: Material benefit**  Benefit = tax break | **IV1: Reputational benefit**  Benefit = public recognition |
| **IV2: Anticipated benefits** | **Dependent variables**  Moral character [Replication]  “Please rate the extent to which the following traits describe Jeff:” (List: Moral, altruistic, sincere, pure, good, nice, immoral, selfish, insincere, impure, bad, mean)  (1 = *Not at all* to 7 = *Extremely)*  [Last 6 traits were reverse coded]  Authentic prosocial motivation [Replication]  “How authentic do you find Jeff's decision to donate to the local soup kitchen?”  (1 = *Not at all authentic* to 7 = *Extremely authentic*)  “How suspicious are you of Jeff’s intentions?”  (1 = *Not at all suspicious* to 7 = *Extremely suspicious*)  “To what extent do you agree with the following statements:”  - “Jeff has a genuine passion for helping the local soup kitchen”  - “Jeff sincerely cares about hungry families in his community”  - “Jeff donated to the local soup kitchen to benefit hungry families in his community”  (1 = *Strongly disagree* to 7 *Strongly agree*)  Inferred emotion type [Extension]  “Please rate the extent to which you think Jeff truly feels the following emotions when thinking about hungry families in his community:” [options: sympathy, distress, sadness, happiness, guilt, anger, pity, discomfort, compassion]  (1 = *Not at all* to 7 *Very much*)  Self-focus [Extension]  “To what extent was Jeff thinking about himself?”  (1 = *Not at all* to 7 *Very much*)  Other-focused [Extension]  “To what extent was Jeff thinking about others”  (1 = *Not at all* to 7 *Very much*)  **Benefit manipulation check [Replication]**  “What benefit does Jeff expect to receive from donating to the soup kitchen?”  (Choices: *Receive a tax break*, *get his name in the newspaper*, *feel happy*, *no benefit*, *I don’t know*)  **Comprehension checks**  With condition matched validation checks.  “What benefit does Jeff [not] expect to receive from donating to the soup kitchen?”  (*Receive a tax break*, *Public recognition*, *feel happy*, *no benefit*, *it doesn’t say*)  “Was Jeff expecting some benefit from donating to the local soup kitchen?”  (*Yes*, *no*, *it doesn’t say*) | | |
|
|
| **IV2: No anticipated benefits** |
|
|

## Mandatory validated comprehension checks

We used comprehension checks to ensure that participants read and understood the instructions and the scenarios, with multiple choice questions that participants had to answer correctly in order for them to proceed to the dependent variables (presented in random order). This is a deviation from the target’s procedure and was meant to ensure that participants carefully read the scenario, and understood and processed the manipulated attributes. We have previously implemented these adjustments in other replications, and found those important to ensure attentiveness and focus on the manipulated attributes. We consider these to be a more conservative test than the target’s design. We chose to use forced validations rather than exclusions, given that even attentive participants may get some comprehension checks wrong, and possible exclusions create forks in analyses adding flexibility in analyses and complexity leading to possible confusion regarding interpretability of results.

In Study 3 we asked the following questions: 1) “What were the person’s galvanic skin responses?” (Low; neutral; high), 2) “What does the galvanic skin response measure?” (Sympathy; distress; it doesn’t say), and 3) “How much did the person donate?” ($0 - Did not donate; Donated $3; it doesn’t say).

In Study 6 we asked the following questions: 1) “Was Jeff expecting some benefit from donating to the local soup kitchen?” (Yes; no; it doesn’t say), and 2) “What type of benefit was mentioned as a possible outcome for donating to the soup kitchen?” (Receive a tax break, public recognition, feel happy, no benefit, it doesn’t say).

## Study 3: Dependent measures

### Replication

#### Moral character

We used the same scales used in the original experiment which were based on existing measures of moral character (e.g., Reeder & Spores, 1983; Wojciszke et al., 1998). Participants indicated the extent to which they thought a person fit 12-character traits (6 positive, 6 negative; = .XX; 1 = *Not at all*; 7 = *Extremely*).

#### Inferred emotion type

We used the same scales used in the target article. Participants indicated the extent to which the person truly felt nine different emotions (adapted from Batson et al., 1983) (1 = *Not at all*; 7 = *Extremely*). Results for “sympathy” and “compassion” were combined to create a measure of perceived empathy ( = .XX), and “distress” and “discomfort” formed a measure of perceived distress ( = .XX). To mirror the original article, results for perceived distress were also used as the distress manipulation check.

#### Emotion level motivation

Participants rated the extent to which they thought the person was motivated by feelings (1 = *Not at all*; 7 = *Extremely*).

#### Likelihood of donating

Participants in the control condition also rated “How likely is it that the person donated money to the African Children’s Fund?” (1 = *Not at all likely*; 7 = *Extremely likely*).

#### Humanness

This measure was not reported in the manuscript but was adopted from the Qualtrics survey which the authors provided. In this measure, participants answered how human they found the person to be (1 = *Not at all human*; 7 = *Extremely human*).

### Extensions

#### Authentic prosocial motivation

We adapted a five-item measure of authentic prosocial motivation ( = .XX) from earlier studies in the same article. The authors noted that this measure was not included in the target article because it only seemed applicable to the donation condition (p. 401). Hence, we made minor adjustments to make it applicable across all three conditions (donated, receive cash, and no information). The questions were as follows: 1) “How genuine do you find the person's intentions?” (1 = *Not at all genuine*, 7 = *Extremely genuine*), 2) “To what extent is the person’s behavior reflective of their intentions?” (1 = *Not at all reflective*, 7 = *Extremely reflective*), and the extent to which they agreed with the following statements: (1 = *Strongly disagree*, 7 = *Strongly agree*) 3) “The person has a genuine passion for the African Children's Fund”, 4) “The person sincerely cares about children in Africa”, and 5) “The person wants to help children in Africa”.

#### Self and other focus

We added a measure which we adapted from earlier studies within the same target article, which examined whether or not the person’s actions were interpreted as self-focused. Participants rated the extent to which they thought the person was thinking about him/herself (1 = *Not at all*; 7 = *Extremely*).

Participants also rated the extent to which they thought the person was thinking about others (1 = *Not at all*; 7 = *Extremely*).

#### Perceived emotional benefits

Participants in the donation condition were rated the extent to which they agreed with the assessment that the person donated to the African Children’s Fund to make him/herself feel better (1 = *Strongly disagree*; 7 = *Strongly agree*). Participants in the cash condition were asked the extent to which they agreed with the statement that the person opted to take the cash to make him/herself feel better. Participants in the control condition were asked the extent to which they agreed with the statement that the person wanted to make him/herself feel better. Responses in the cash and control conditions were reverse coded.

## Study 6: Dependent measures

### Replication

The measures used to assess **moral character** ( = .XX) and perceived emotion type were the same as Study 3 (**perceived distress**: = .XX, **perceived empathy**: = .XX).

#### Authentic prosocial motivation

We adopted the scale used in the target article and only modified the subject of the statement such that it referred to the person as “Jeff” instead of “the donor”. Participants had to answer the following questions ( = .XX): 1) “How authentic do you find Jeff's decision to donate to the local soup kitchen?” (1 = *Not at all authentic*, 7 = *Extremely authentic*), 2) “How suspicious are you of the person's intentions?” (1 = *Not at all suspicious*, 7 = *Extremely suspicious*), and “To what extent do you agree with the following statement[s]” (1 = *Strongly disagree*, 7 = *Strongly agree*): 3) “Jeff has a genuine passion for helping the local soup kitchen”, 4) “Jeff sincerely cares about hungry families in his community”, and 5) “Jeff donated to the local soup kitchen to benefit hungry families in the community”.

### Extensions

#### Self and other focus

The measures used were the same as Study 3, except for the other focus measure which was modified so that it referred to ‘Jeff’ rather than to ‘the person’.

## Deviations

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

###### Table 6 *Adjustments to the target article’s Studies 3 and 6 methods and design*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Study** | **Factor** | **Target article** | **Adjustment in current study** | **Reason for change** |
| 3 | Study design | Participants in the no donate/gift-card condition were told that the targets opted to receive the bonus $3 iTunes gift card | Instead of a $2 iTunes gift card, the rewards were changed to $3 in cash | An iTunes gift card is not as versatile and therefore might not be as ‘alluring’ as receiving the reward in cash |
|  | Stimulus | Participants were shown the target’s GSR records and “Donation summary” (p. 401) | i) Added “Amount received” for participants in the cash condition  ii) Added a red box around “Donated?” and “Amount donated”/ “Amount received” | i) Lack of information on how the “Donation summary” looked like  ii) Ensure that participants focused on the manipulation in the stimulus |
| 6 | Stimulus | Conflation of outcome and expectations | We adjusted to only manipulate expectations and not vary outcomes (e.g., expectation of reputation conflated with an anonymous or public donation) | Addressing a confound with only a minor adjustment, aiming for better alignment with theory. |
| 3 and 6 | Procedure | Comprehension, manipulation and attention checks were combined with the DVs | Forced comprehension with validation. Participants are first shown the stimulus, then asked the comprehension and manipulation checks. They could only proceed to the measures once they have responded correctly to the questions. | Ensured participants read, understood, and processed the scenario and manipulation before responding to the questions. We consider this a more conservative test than the one in the target article. |

*Note:* Materials provided by the authors of the original article indicated that the stimulus used in Study 3 was $3/$3 iTunes gift card/No information rather than $2/$2 iTunes gift card/No information as written in the article.

## Evaluation criteria for replication findings

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

We pre-registered our overall strategy to conclude an overall successful replication of the target article if both studies showed a signal in the same direction as the target article, a failed replication if neither study showed a signal in the same direction as the original, and mixed findings if only one study showed a signal in the same direction as the original. In both studies we would consider it a successful replication only if both the main effect and the interactions are supported, and if only the main effect is supported, then that would be considered mixed findings.

## 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** |
| Effect/hypothesis | Same |  |
| IV construct | Similar | Study 3’s gift-card condition was changed into a cash condition.  Study 6’s ‘did not expect X benefit’ conditions were adjusted as not to conflate expected outcomes |
| DV construct | Similar |  |
| IV operationalization | Same |  |
| DV operationalization | Same |  |
| IV stimuli | Similar | Stimulus was based on materials provided by the original authors (Study 2) and screenshots in the Appendix (Study 5) of the target article as the stimuli was similar to the ones used for Studies 3 and 6. The stimuli in Study 6 was slightly adjusted to address expectations-outcome confound. |
| DV stimuli | Same |  |
| Procedural details | Similar | Forced validated comprehension checks were added before the measures. |
| Physical settings | Similar | Both were conducted online using Qualtrics and recruited participants through Prolific. |
| Contextual variables | Different | Original studies were conducted in 2014 or before, the replication was conducted in 2024. The time gap may cause variations. |
| Population (e.g., age) | Different | 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

All the analyses were conducted using a combination of JAMOVI(Version 2.2.2) and R (Version 4.3.2.). We wrote our planned analysis code (see OSF folder) using packages “pwr”, “effectsize” (Version 0.8.6), “MOTE” (Version 1.0.2; Buchanan et al., 2022), “MBESS” (Version 4.9.3; Kelley et al., 2018), “ggstatsplot” (Version 0.12.2; Patil, 2021), “psych” (Version 2.4.1), “rstatix” (Version 0.7.2), “afex” (Version 1.3-0), “car” (Version 3.1-2), “jmv” (Version 2.4.11), “haven” (Version 2.5.4), “emmeans” (Version 1.10.0), “rstantools” (Version 2.4.0). We note that we applied the Tukey p-values adjustment to all analyses that include post-hoc comparisons.

### Replication: Two-way ANOVA

***Study 3***

To mirror the target article’s analyses, we first ran two-way ANOVAs (2 × 3) to examine the effect of distress and chosen action on perceptions of moral character, inferred emotion types, emotion level motivation, inferred emotion type, and humanness.

***Study 6***

To mirror the target article’s analyses, we ran two-way ANOVAs (2 × 3) to assess the effect of expected benefit and benefit type on perceptions of moral character, authentic prosocial motivation, inferred emotion, and humanness.

### Replication: Additional analyses

#### Study 3

For participants in the control condition, we also ran independent samples t-test on the likelihood of donating which compared answers from the high distress condition with that in the low distress condition.

### Extensions: Two-way ANOVAs

To analyze our extensions, we also conducted two-way ANOVAs (2 × 3) to assess the effect of distress level and chosen action on authentic prosocial motivation, self-focus, other-focus, and emotion benefits motivation. Similarly, in Study 6 we ran two-way ANOVAs (2 × 3) to examine the effect of expected benefit and benefit type on self-focus and other-focus. The alpha value is set to **.005** for all the extensions.

### 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 .**001**. Our planned sample size is large enough to provide sufficient statistical power to conduct moderation analyses and examine the order effects, if needed.

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

### Outliers and exclusions

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

### Bayesian analyses

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

# 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 findings in comparison to the target article in Table 8. In the sections below we provide more detail for each study and analysis.

###### Table 8 *Replication of Studies 3 and 6: Summary of statistical tests, effects, and evaluation of findings*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Our replication** | | | **Target article findings** | | |  |
| **S** | **DV** | **IV** | ***F* statistic** | ***p*** | ***η2p* and CI** | ***F* statistic** | ***p*** | ***η2p* and CI** | **Interpretation** |
| 3 | Distress manipulation check | Distress | *F*(1, 994) = 0.61 | .434 | 0.00  [0.00, 0.02] | *F*(1, 464) = 432.68 | <.001 | 0.48  [0.42, 0.54] | No signal; inconsistent; smaller |
| Distress x action | *F*(2, 994) = 0.11 | .893 | 0.00  [0.00, 0.01] | *F*(2,464) = 3.04 | .05 | 0.01  [0.00, 0.04] | No signal; inconsistent; smaller |
| Moral character | Distress | *F*(1, 994) = 1.73 | .189 | 0.00  [0.00, 0.02] | *F*(1, 464) = 51.56 | <.001 | 0.10  [0.05, 0.15] | No signal; inconsistent; smaller |
| Distress x action | *F*(2, 994) = 0.70 | .147 | 0.00  [0.00, 0.02] | *F* (2,464) = 12.73 | <.001 | 0.05  [0.02, 0.09] | No signal; inconsistent; smaller |
|  | **DV** | **IV** | ***F* statistic** | ***p*** | ***η2p* and CI** | ***F* statistic** | ***p*** | ***η2p* and CI** | **Interpretation** |
| 6 | Moral character | Expected benefit | *F*(1, 994) = 0.56 | .454 | 0.00  [0.00, 0.02] | *F*(1, 280) = 34.68 | <.001 | 0.11  [0.05, 0.18] | No signal; inconsistent; smaller |
| Expected x benefit type | *F*(2, 994) = 0.13 | .880 | 0.00  [0.00, 0.01] | *F*(1,280)=25.43 | <.001 | 0.15  [0.08, 0.23] | No signal; inconsistent; smaller |
| Authentic prosocial motivation | Expected benefit | *F*(1, 994) = 0.31 | .578 | 0.00  [0.00, 0.01] | *F*(1, 280) = 68.59 | <.001 | 0.20  [0.12, 0.28] | No signal; inconsistent; smaller |
| Benefit type | *F*(2, 994) = 2.63 | .072 | .01  [0.00, 0.01] | *F*(1,280) = 7.46 | .001 | 0.03  [0.00, 0.07] | No signal; inconsistent; smaller |
| Expected x benefit type | *F*(2, 994) = 0.07 | .933 | .00  [0.00, 0.00] | F(1,280) = 25.82 | <.001 | 0.16  [0.08, 0.23] | No signal; inconsistent; smaller |

*Note*. All analyses listed in the table were ANOVAs, *CI* = 95% confidence intervals. The interpretation of outcome was based on LeBel et al. (2019).

## Study 3

We summarized the descriptives in Table 9, the statistical tests in Table 10, and the correlations between the measures in Table 11.

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Cash | | Donate | | No information | |
|  | High distress (*n* = 166) | Low distress (*n* = 167) | High distress (*n* = 167) | Low distress (*n* = 167) | High distress (*n* = 166) | Low distress (*n* = 167) |
| Moral character | 3.88 (0.58) | 4.03 (0.60) | 4.02 (0.58) | 3.99 (0.64) | 3.92 (0.62) | 3.94 (0.59) |
| Perceived distress | 4.16 (1.54) | 4.15 (1.43) | 4.19 (1.29) | 4.08 (1.41) | 4.07 (1.44) | 3.97 (1.41) |
| Emotion benefits motivation (Extension) | 3.75 (1.96) | 3.88 (1.87) | 3.983 (1.90) | 3.93 (2.01) | 4.26 (2.05) | 3.95 (1.99) |
| Authentic prosocial motivation (Extension) | 4.00 (0.84) | 4.02 (0.90) | 4.00 (0.92) | 3.92 (0.88) | 4.05 (0.97) | 4.00 (0.92) |
| Self-focus (Extension) | 3.92 (2.05) | 4.02 (2.04) | 4.01 (2.06) | 3.78 (1.95) | 3.84 (1.98) | 4.00(2.03) |
| Other-focus (Extension) | 3.61 (1.92) | 4.15 (1.86) | 3.93 (2.09) | 4.21 (1.88) | 3.93 (1.97) | 4.05 (2.00) |

*Note*: Format = mean (standard deviations).

###### 

###### Table 10 *Study 3: Summary of statistical tests and effects*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Measures** | **Independent variables** | ***F*-statistic** | ***p*** | ***η2p*** |
| Moral character | Distress | *F*(1, 994) = 1.73 | .189 | .003 |
| Action | *F*(2, 994) = 1.31 | .271 | .002 |
| Distress × action | *F*(2, 994) = 0.696 | .147 | .004 |
| Distress manipulation | Distress | *F*(1, 994) = 0.611 | .434 | .000 |
| Action | *F*(2, 994) = 0.874 | .417 | .002 |
| Distress × action | *F*(2, 994) = 0.113 | .893 | .000 |
| Authentic prosocial motivation (Extension) | Distress | *F*(1, 994) = 0.312 | .577 | .000 |
| Action | *F*(2, 994) = 0.370 | .691 | .001 |
| Distress × action | *F*(2, 994) = 0.402 | .669 | .001 |
| Emotion benefits motivation (Extension) | Distress | *F*(1, 994) = 0.396 | .529 | .000 |
| Action | *F*(1, 994) = 1.83 | .162 | .004 |
| Distress × action | *F*(1, 994) = 1.04 | .353 | .002 |
| Self-focus (Extension) | Distress | *F*(1, 994) = 0.011 | .915 | .000 |
| Action | *F*(2, 994) = 0.114 | .894 | .000 |
| Distress × action | *F*(2, 994) = 0.917 | .400 | .002 |
| Other-focus (Extension) | Distress | *F*(1, 994) = 6.29 | .012 | .006 |
| Action | *F*(2, 994) = .077 | .465 | .002 |
| Distress × action | *F*(2, 994) = .376 | .376 | .002 |

*Note*. \* *p* < .05; \*\* *p* < .01; \*\*\* *p* < .001.

###### Table 11 *Study 3: Correlations between main measures*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | *α* | 1 | 2 | 3 | 4 | 5 |
| 1 - Moral character | .09 | - |  |  |  |  |
|  |  |  |  |  |  |
| 2 - Perceived Distress | .03 | -.01 | - |  |  |  |
|  | [-.07, .05]  (-.02) |  |  |  |  |
| 3 - Authentic prosocial motivation | .01 | .01 | .00 | - |  |  |
|  | [-.05, .07]  (.01) | [-.07, .05]  (.00) |  |  |  |
| 4 - Emotional benefits motivation | - | -0.04 | -.04 | .03 | - |  |
|  | [-.1, .02]  (-.04) | [-.10, .02]  (.02) | [-.03, .09]  (-.05) |  |  |
| 5 - Self-focus | - | .00 | .02 | .03 | -.02 | - |
|  | [-.06, .07]  (.01) | [-.04, .08]  (.02) | [-.03, .09]  (.03) | [-.08, .0.39]  (-.02) |  |
| 6 - Other-focus | - | -.02 | .01 | -.05 | -.07 | -.01 |
|  | [-.04, .08]  (-.02) | [-.05, .07]  (.01) | [-.11, .02]  (-.05) | [-.15, .00]  (-.07) | [-.07, .05]  (-.01) |

*Note*. (*n/N* = XXX). Format: Pearson’s correlations [confidence intervals] (Spearman’s rho).   
\* *p* < .05; \*\* *p* < .01; \*\*\* *p* < .001.

### Replication

#### Distress manipulation check

We conducted a two-way ANOVA and found no support for a main effect of manipulated distress on perceived distress, *F*(1, 994) = 0.61, *p* = .435, *η2p* = .001, 95% CI [0.00, 0.01], for a main effect of action on perceived distress, *F*(2, 994) = 0.87, *p* = .417, *η2p* = .00, 95% CI [0.00, 0.01], nor for an interaction between distress levels and chosen action *F*(2, 994) = 0.11, *p* = .893, *η2p* = .000, 95% CI [0.00, 0.00]. We found no support for manipulated distress or actions targets took as affecting perceived distress levels. In comparison, Barasch et al. (2014) found support for an effect of manipulated distress on perceived distress *F*(1, 464) = 432.68, *p* <.001, *η2p* = .48, 95%, CI [0.42, 0.54], and for an interaction effect between distress and action, *F*(1, 464) = 3.04, *p* = .05, *η2p* = 0.01, 95% CI [0.00, 0.04].

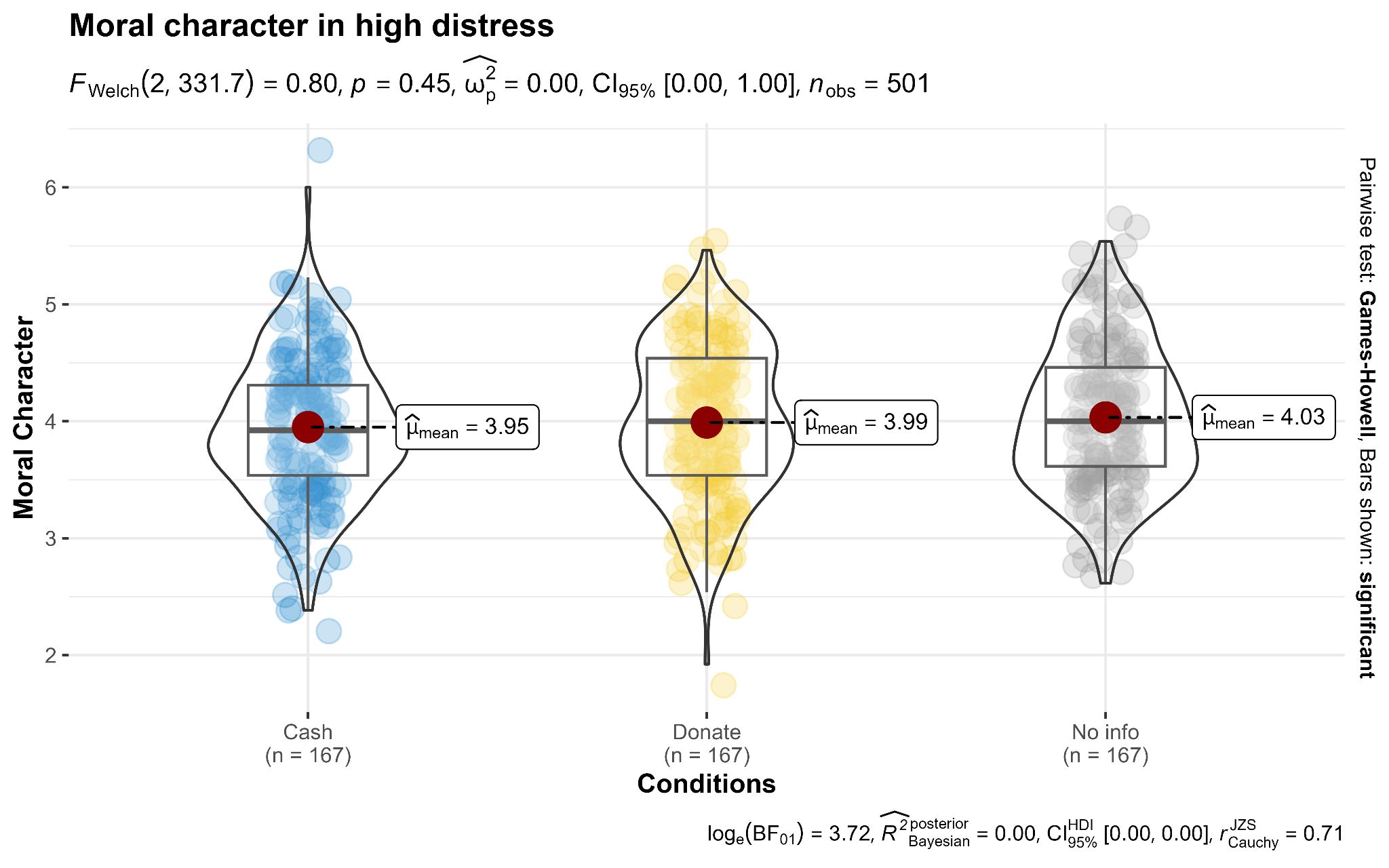
#### 

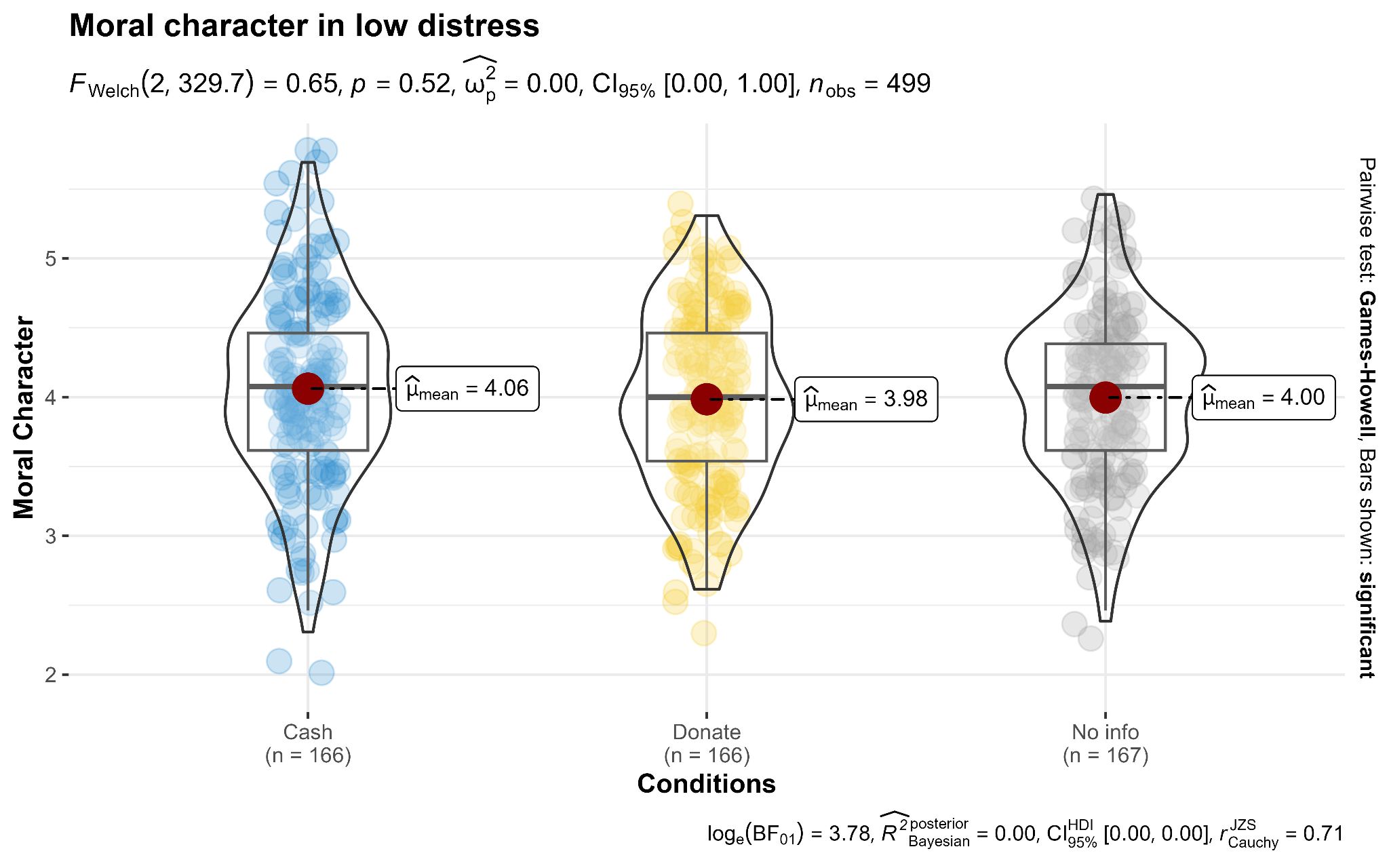
#### Moral character

We conducted a two-way ANOVA and found no support for a main effect of distress level on perceived moral character, *F*(1, 994) = 1.73, *p* = .189, *η2p* = .001, 95% CI [0.00, 0.01], a main effect of chosen action on perceptions of moral character, *F*(2, 994) = 1.31, *p* = .270, *η2p* = .003, 95% CI [0.00, 0.01], nor for an interaction between distress level and chosen action, *F*(2, 994) = 1.92, *p* =.147, *η2p* = .004, 95% CI [0.00, 0.01]. We found no support for Hypotheses 1a and 2 that the more distressed a donor feels regarding harm to others, the more likely they are to be judged as being more moral than others. In comparison, the target article found support for an effect of distress (*F*(1, 464) = 51.56, *p* <.001, *η2p* = 0.10, 95% CI [0.05, 0.15]), action (*F*(1, 464) = 105.69, *p* <. 001, *η2p* = = 0.31, 95% CI [0.25, 0.37]), and a distress by action interaction (*F*(1, 464) = 12.73, *p* <.001, *η2p* = 0.05, 95% CI [0.02, 0.09]) on moral character. We provided a summary plot in Figure 1.

###### 

###### Figure 1 *Study 3: Moral character*





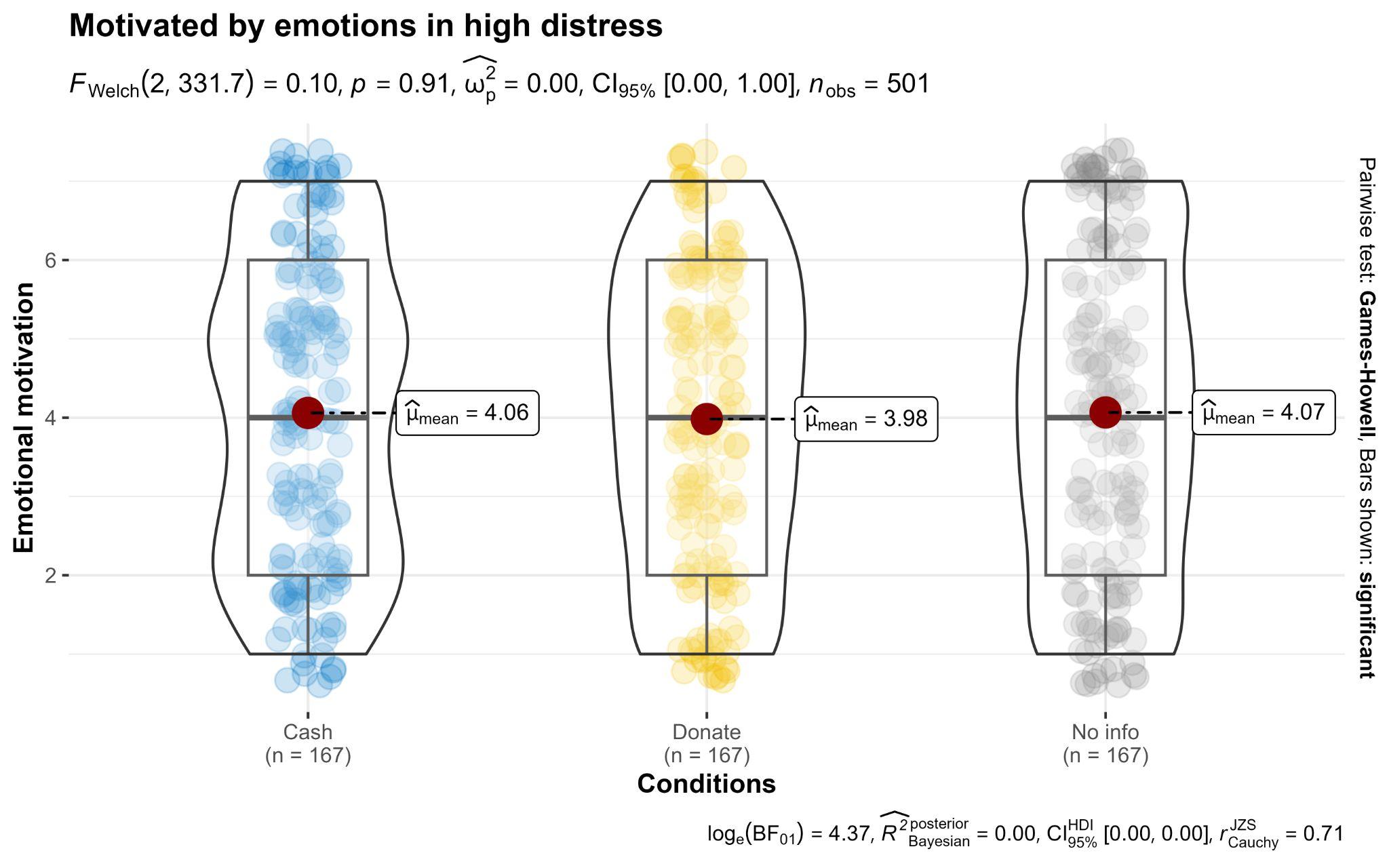
*Note*. Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

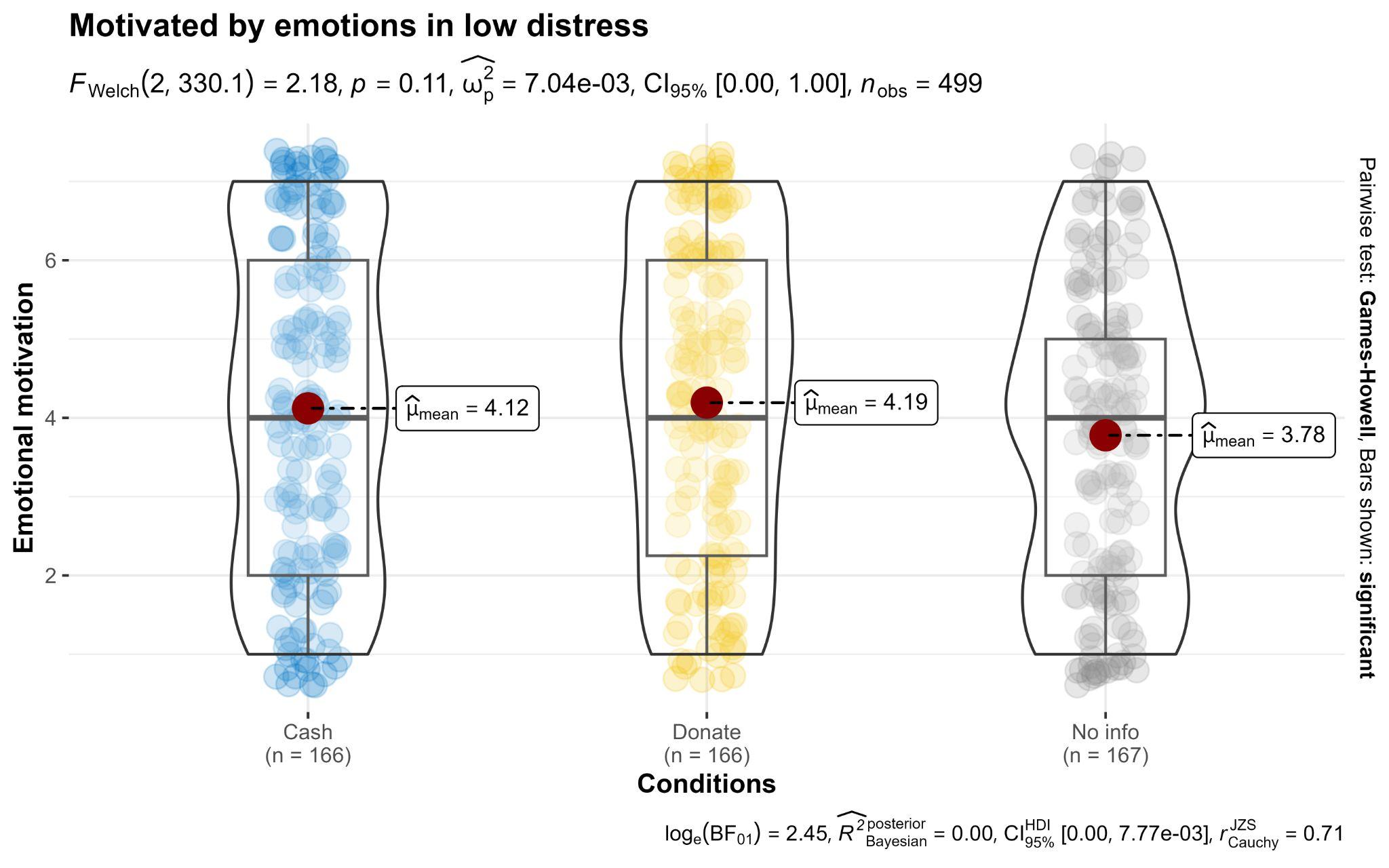
#### 

#### Emotional motivation

We conducted a two-way ANOVA and found no support for a main effect of manipulated distress on perceived emotional motivation *F*(1, 994) = 0.813, *p* = .368, *η2p* = .00, 95% CI [0.00, 0.01], a main effect of chosen action on perceived emotional motivation *F*(2, 994) = 0.917, *p* = .400, *η2p* = .002, 95% CI [0.00, 0.01], or an interaction effect observed between manipulated distress and the chosen action *F*(2, 994) = 4.98, *p* = .007, *η2p* = .0001, 95% CI [0.00, 0.02]. The target article did not report their findings for this measure, it is likely that this was seen as a non-essential measure. We provided a summary plot in Figure 2.

###### Figure 2 *Study 3: Perceived emotional motivation*





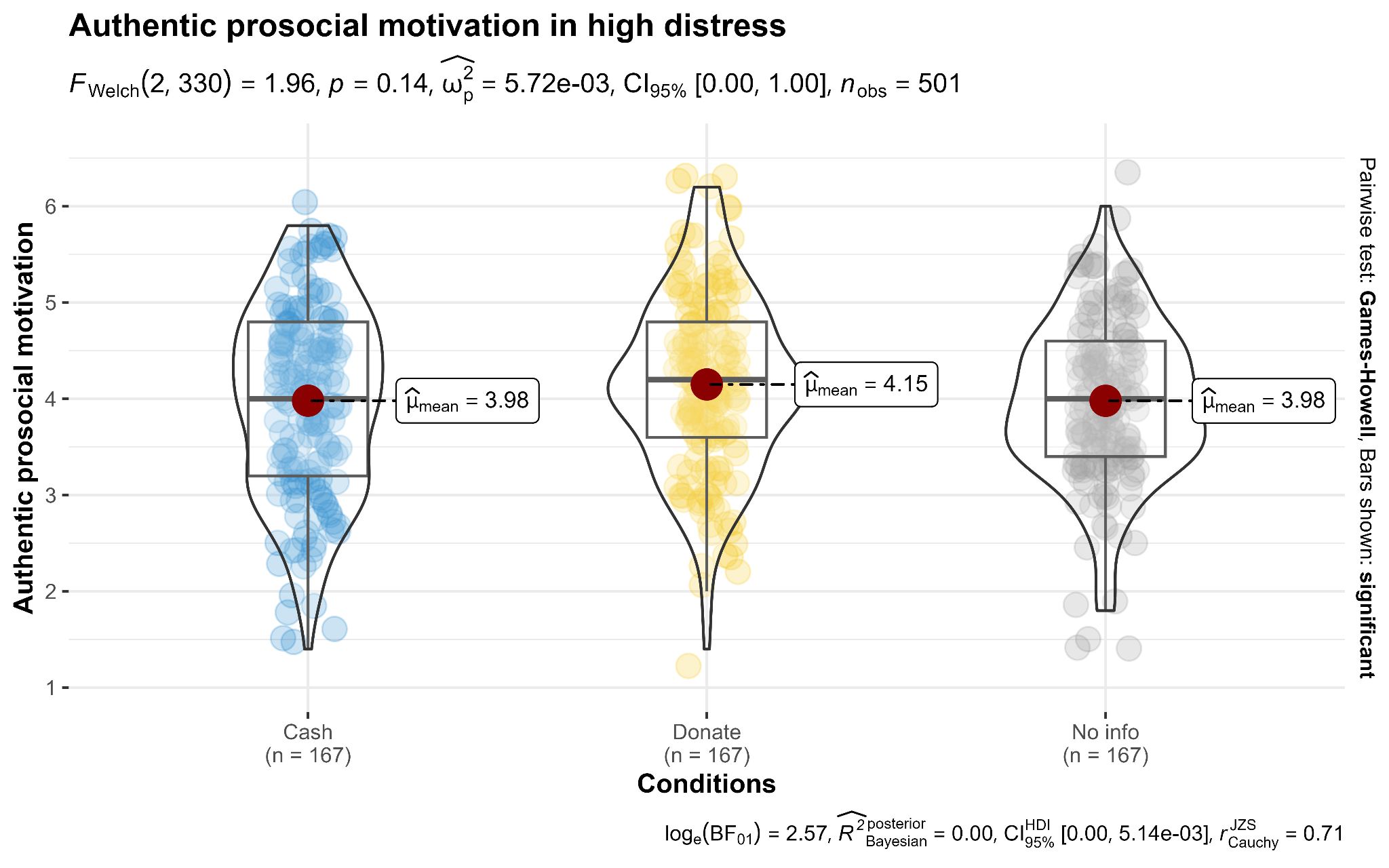
*Note*. Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

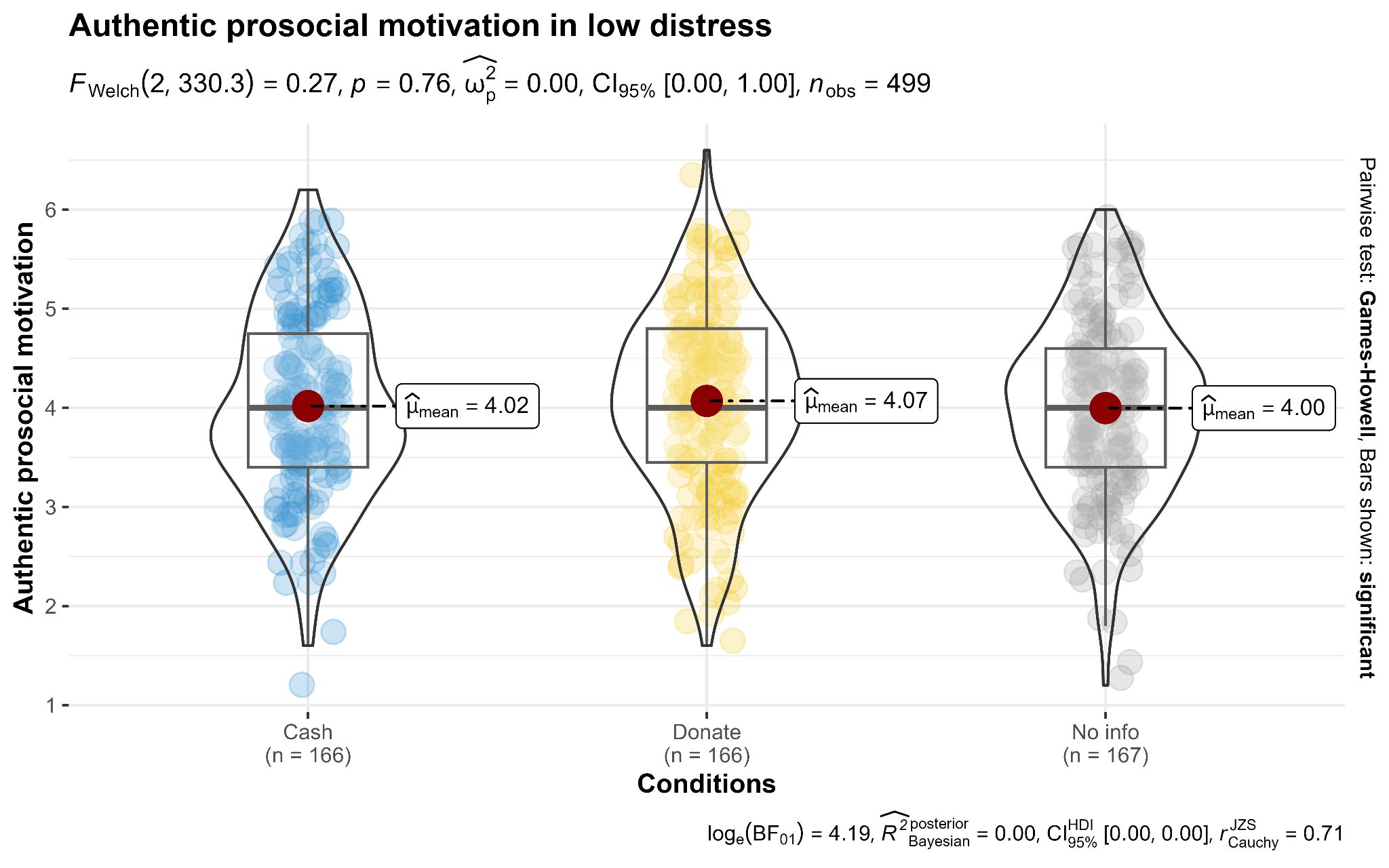
### Extensions

#### Authentic prosocial motivation

We found no support for a main effect of distress level on authentic prosocial motivation (Hypothesis 1b; *F*(1, 994) = .312, *p* = .577, *η2p* = .000, 95% CI [0.00, 0.01]), a main effect of action on authentic prosocial motivation (*F*(2, 994) = .370, *p*  = .691, *η2p* = .00, 95% CI [0.00, 0.01]), nor an interaction between distress level and action on authentic prosocial motivation (*F*(2, 994) = .402, *p* = .669, *η2p* = .000, 95% CI [0.00, 0.01]). We provided a summary plot in Figure 3.

###### Figure 3 *Study 3: Perceived authentic prosocial motivation*



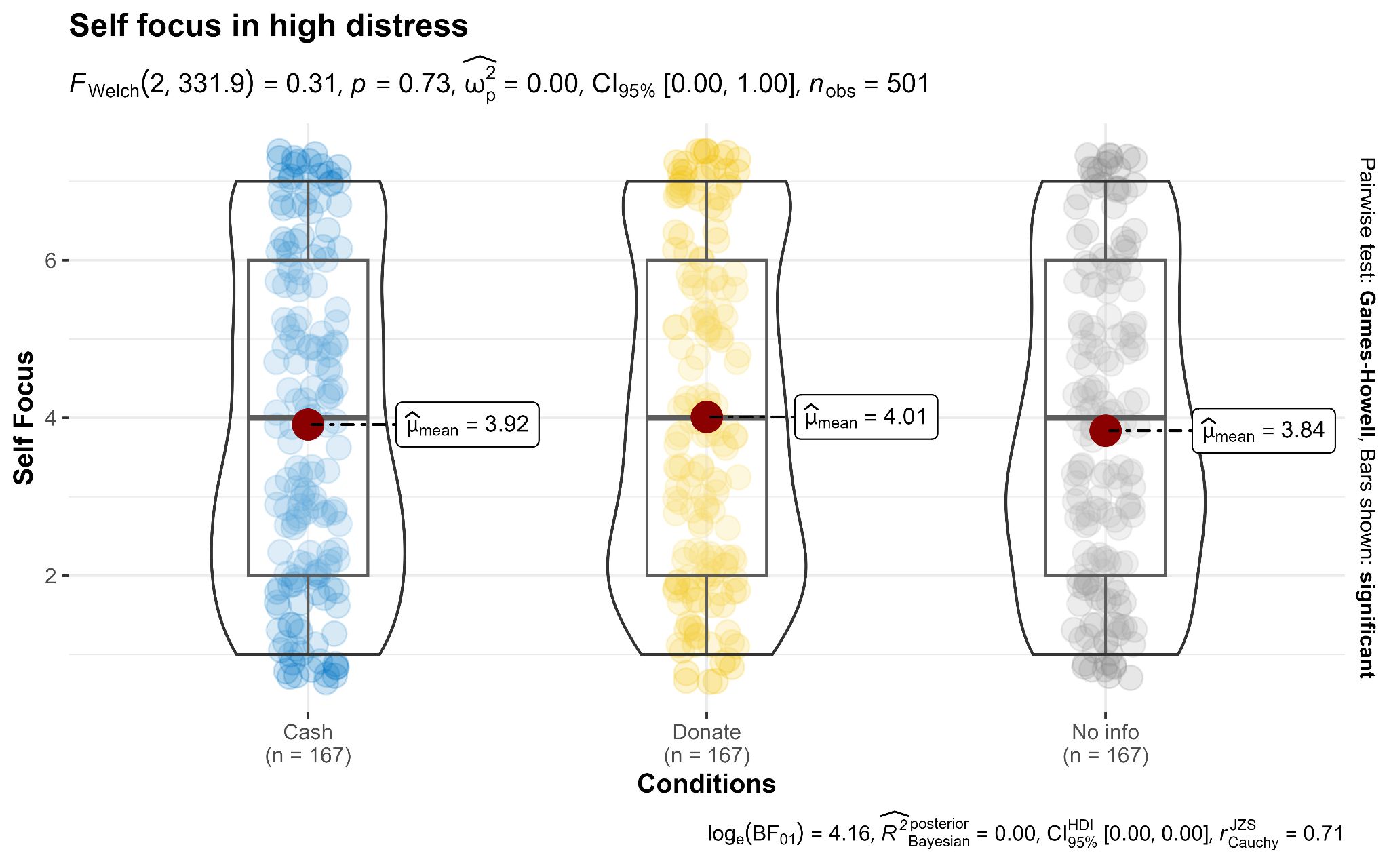


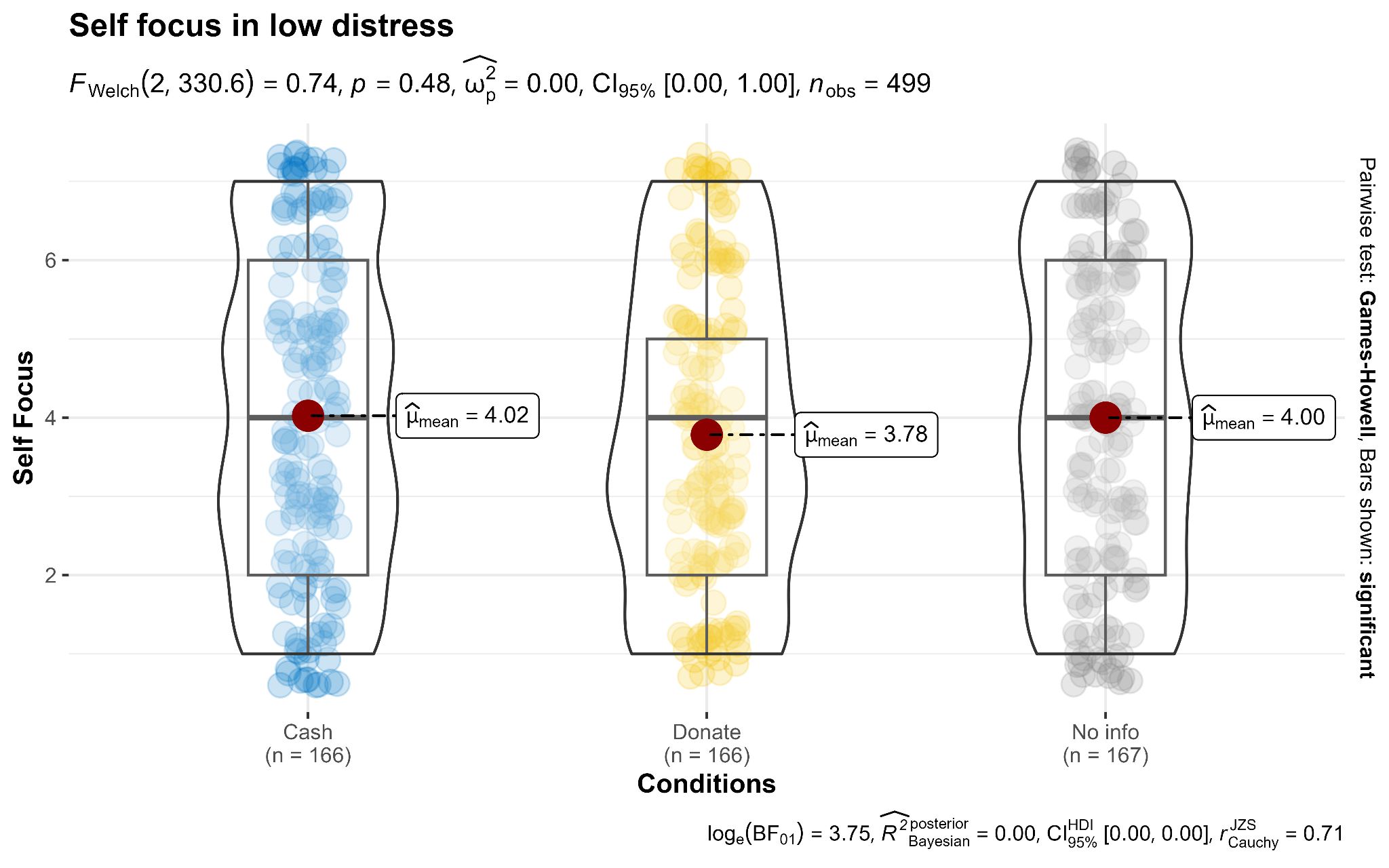
*Note*. Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

#### Self-focus

We conducted a two-way ANOVA found no support for a main effect of distress level on perceived self-focus, *F*(1, 994) = 1.59, *p* = .207, *η2p* = .00, 95% CI [0.00, 0.01], a main effect of action on perceived self-focus, *F*(2, 994) = .05, *p* = .955, *η2p* = .000, 95% CI [0.00, 0.00], or an interaction between distress level and action on perceived self-focus *F*(2, 994) = 1.45, *p* = .235, *η2p* = .00, 95% CI [0.00, 0.01]. We provided a summary plot in Figure 4.

###### Figure 4 *Study 3: Perceived self-focus*



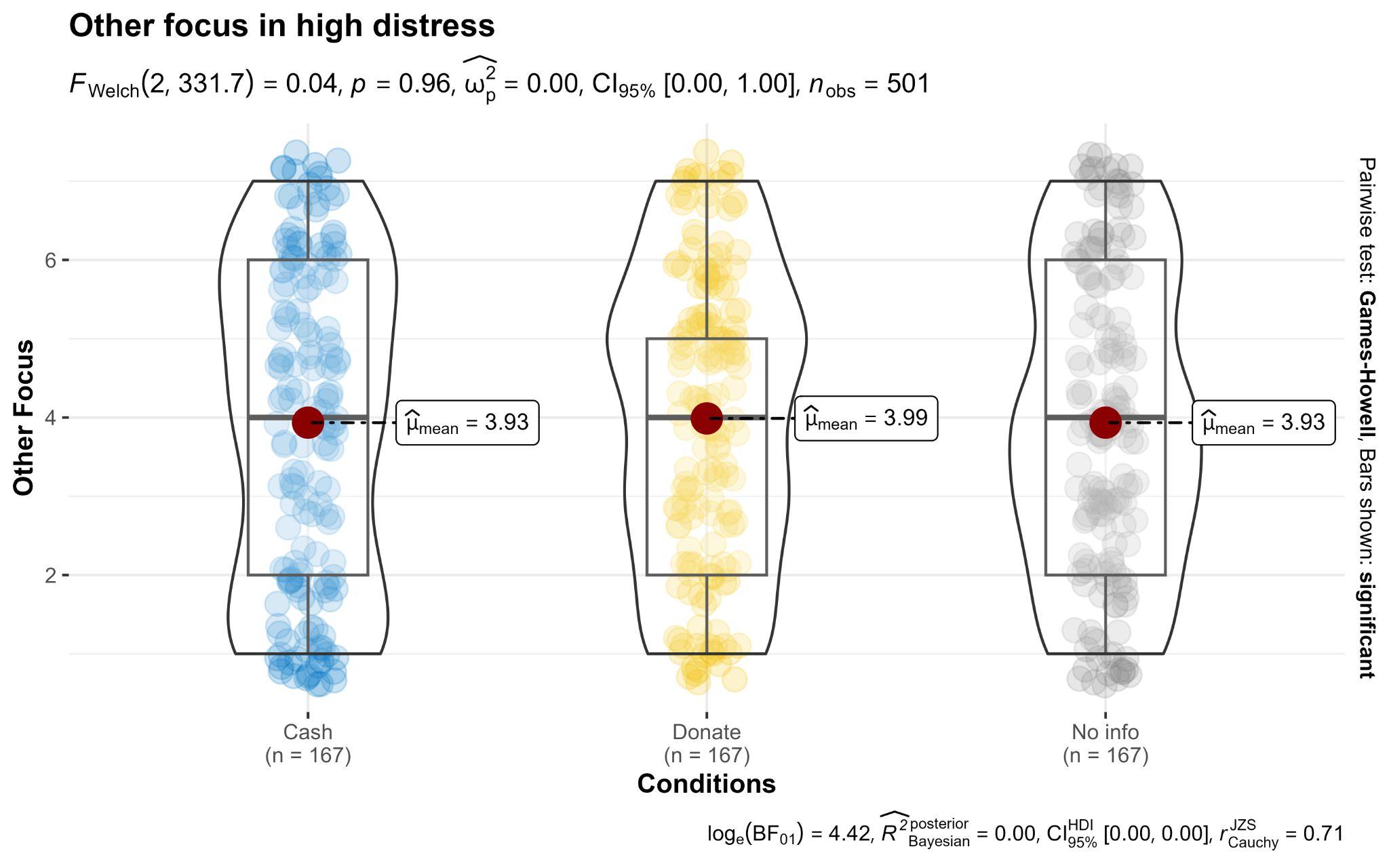


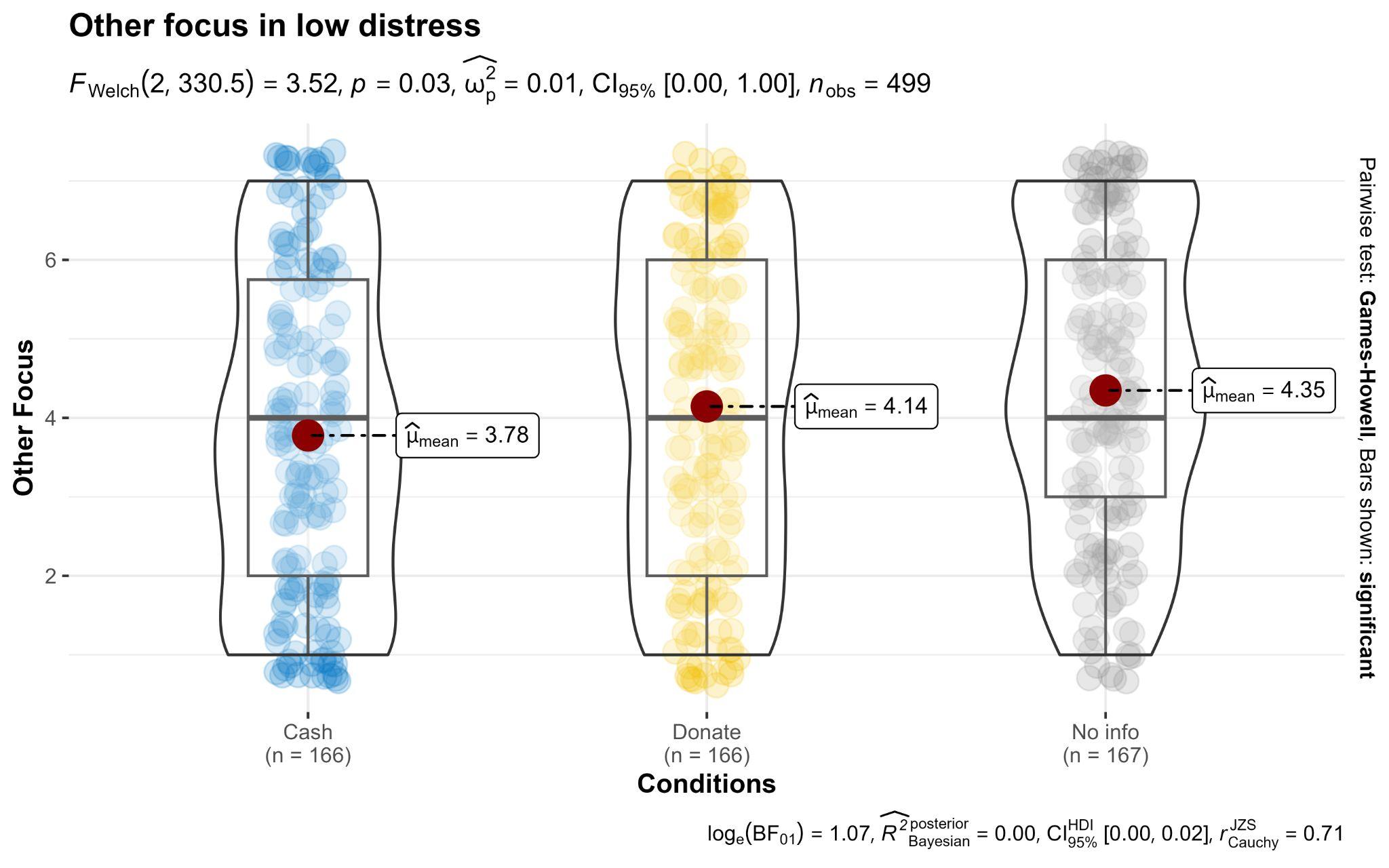
*Note*. Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

#### Other-focus

We conducted a two-way ANOVA and found that there was support for a main effect of distress level on perceived other-focus (*F*(1, 994) = 6.29, *p* = .012, *η2p* = .006, 95% CI [0.00, 0.02]). Specifically, participants perceived targets who felt low distress as being more other-focused than participants who felt high distress (*t*(994) = 2.51, *pTukey* = .012, *d* = .16, 95% CI [0.03, 0.28]). However, we found no support for Hypothesis 3 in that we found no support for the target’s action impacting perceived other-focus (*F*(2, 994) = .767, *p* = .464, *η2p* = .002, 95% CI [0.00, 0.01]), or for an interaction between distress and chosen action on perceived other focus (*F*(2, 994) = .979, *p* = .376, *η2p* = .002, 95% CI [0.00, 0.01]). We provided a summary plot in Figure 5.

###### Figure 5 *Study 3: Perceived other-focus*

**

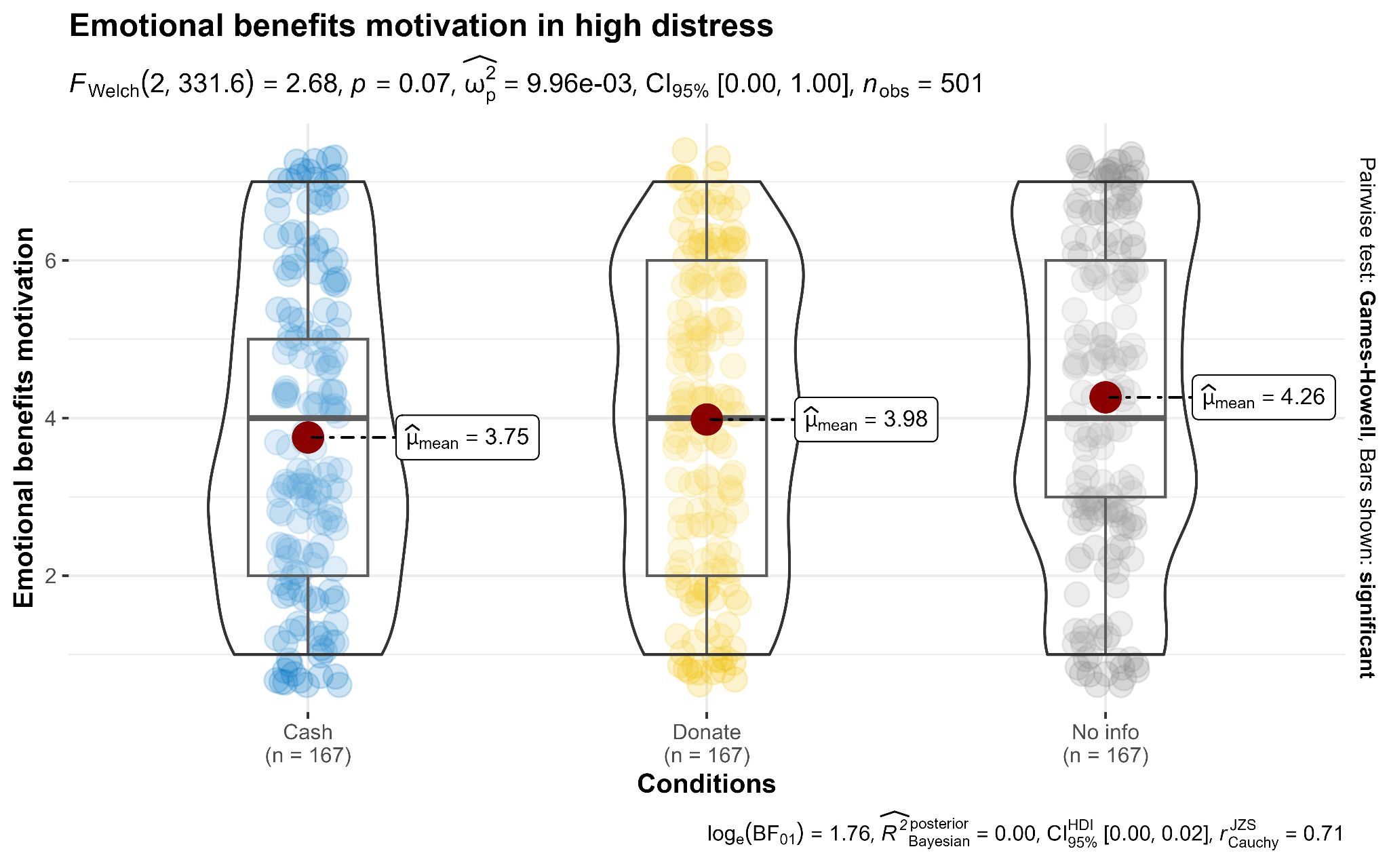
**

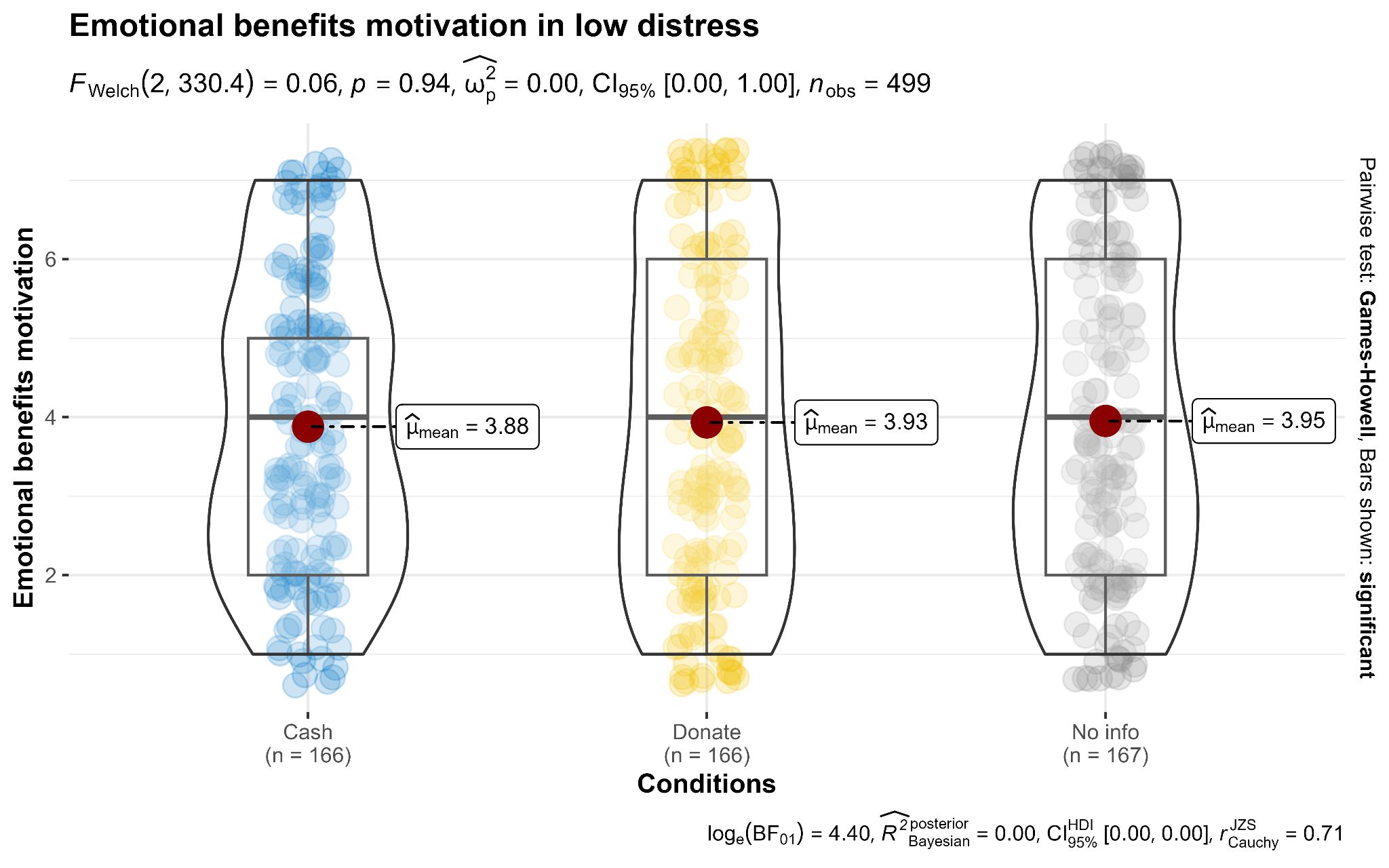
*Note*. Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

#### Emotional benefits motivation

We conducted a two-way ANOVA and found no support for a main effect of distress level on perceived motivation to receive emotional benefits, *F*(1, 994) = .396, *p* = .529, *η2p* = .00, 95% CI [0.00, 0.01], chosen action*, F*(2, 994) = 1.83, *p* = .162, *η2p* = .004, nor was there an interaction found between distress levels and chosen action on perceived motivation to receive emotional rewards, *F*(2, 994) = 1.04, *p* = .353, *η2p* = .002, We provided a summary plot in Figure 6.

###### Figure 6 *Study 3: Emotional benefits motivation*

**

**

*Note*. Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

#### Likelihood of donating

We found no support for differences in donation likelihood between targets experiencing high distress (*M* = 3.92 , *SD* = 2.09) and those experiencing low distress (*M* = 4.00, *SD* = 2.02), *tWelch’s*(330) = -0.35, *p* = .729, *d* = 0.04, 95% CI [-0.25, 0.18]. This was not reported in the original study however we were able to run our own analysis as the original authors generously provided us with the datasets. They found support for differences between distress levels and the likelihood of donating. Specifically, targets in the high distress condition (*M* = 4.50, *SD* = 2.26) were seen as more likely to donate compared to targets experiencing low distress (*M* = 3.02, *SD* = 2.00), *t*(468) = 7.32, *p* = <.001, *d* = .68, 95% CI [0.49, 0.86].

## Study 6

We summarized the descriptives in Table 12, the replication findings in Table 13, and correlations in Table 14.

###### Table 12 *Study 6: Descriptive statistics*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Reputational | | Tax Break | | Emotional | |
|  | Expected (*n* = 166) | Not expected (*n* = 167) | Expected (*n* = 167) | Not expected (*n* = 166) | Expected  (*n* = 167) | Not expected (*n* = 167) |
| Moral character | 4.01 (0.57) | 4.01 (0.68) | 4.01 (0.67) | 3.96 (0.59) | 4.00 (0.63) | 3.96 (0.60) |
| Authentic prosocial motivation | 4.02 (0.95) | 4.07 (0.91) | 3.86 (1.00) | 3.91 (0.90) | 3.99 (0.92) | 4.00 (0.89) |
| Self-focus (Extension) | 4.05 (2.10) | 4.20 (2.08) | 4.13 (1.99) | 4.27 (2.06) | 3.92 (1.95) | 3.79 (2.01) |
| Other-focus (Extension) | 4.08 (1.98) | 4.02 (2.02) | 3.95 (2.01) | 4.01 (1.93) | 4.17 (1.99) | 4.05 (2.05) |

*Note*: The statistics outside the parentheses are the means while the numbers inside the parentheses are the corresponding standard deviations.

###### Table 13 *Study 6: Summary of statistics and effects*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Measures** | **Independent variables** | ***F*-statistic** | ***p*** | ***η2p*** |
| Moral character | Expected | *F*(1, 994) = 0.561 | .454 | .000 |
| Benefit type | *F*(2, 994) = 0.321 | .726 | .001 |
| Expected × benefit | *F*(2, 994) = 0.128 | .880 | .000 |
| Authentic prosocial motivation | Expected | *F*(1, 994) = 0.310 | .578 | .000 |
| Benefit type | *F*(2, 994) = 2.63 | .072 | .005 |
| Expected × benefit | *F*(2, 994) = 0.069 | .933 | .000 |
| Self-focus (Extension) | Expected | *F*(1, 994) = 0.19 | .666 | .000 |
| Benefit type | *F*(2, 994) = 2.70 | .068 | .005 |
| Expected × benefit | *F*(2, 994) = 0.502 | .606 | .001 |
| Other-focus (Extension) | Expected | *F*(1, 994) = 0.081 | .775 | .000 |
| Benefit type | *F*(2, 994) = 0.347 | .707 | .001 |
| Expected × benefit | *F*(2, 994) = 0.163 | .850 | .000 |

###### Table 14 *Study 6: Correlations between main measures*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | *α* | 1 | 2 | 3 |
| 1 - Moral character | -.01 | - |  |  |
|  |  |  |  |
| 2 - Authentic prosocial motivation | .06 | .00 | - |  |
|  | [-.06, .07]  (.00) |  |  |
| 3 - Self-focus | - | .04 | .03 | - |
|  | [-.02, .10]  x(.01) | [-.04, .09]  (-.03) |  |
| 4 - Other-focus | - | .04 | .03 | .01 |
|  | [-.04, .09]  (.02) | [-.02, .10]  (.05) | [-.06, .07]  (.02) |

*Note*. (*n/N* = XXX). Format: Pearson’s correlations [confidence intervals] (Spearman’s rho).   
\* *p* < .05; \*\* *p* < .01; \*\*\* *p* < .001.

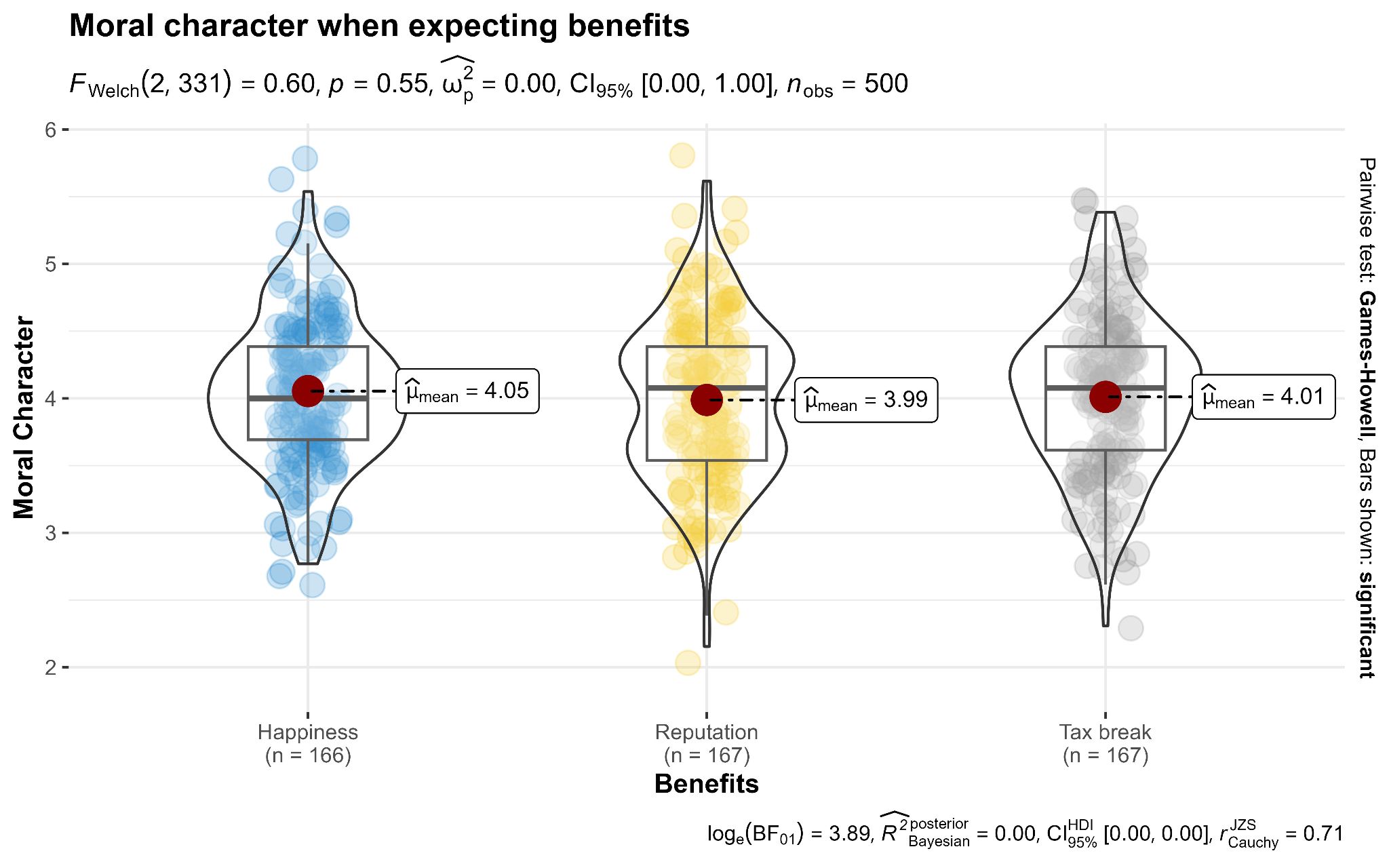
### 

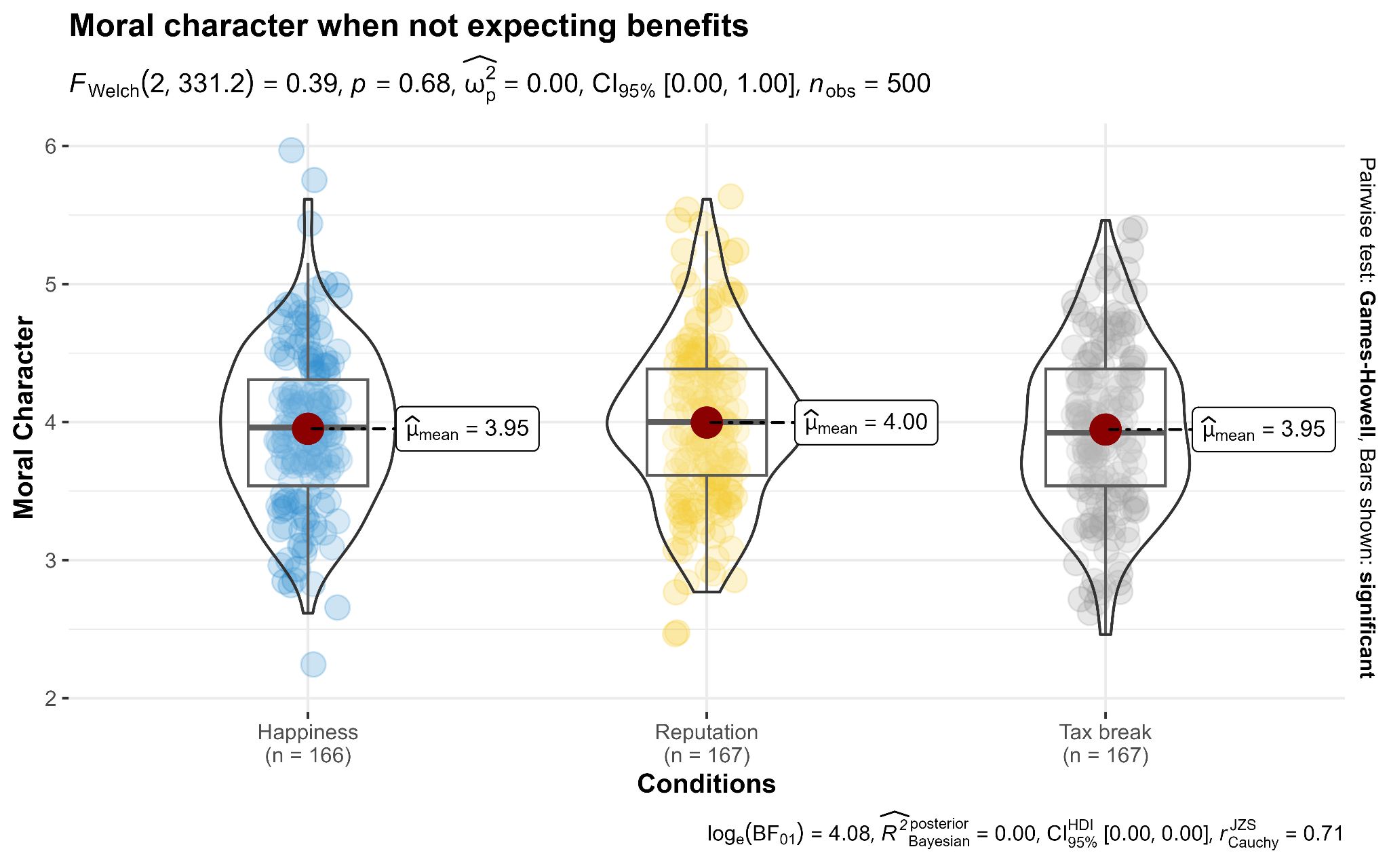
### Replication

#### Moral character

A two-way ANOVA found no support for expecting a benefit on moral character, *F*(1, 994) = .561, p = .454, *η2p* = .000, 95% CI [0.00, 0.01], benefit type on moral character, *F*(2, 994) = .321, *p* = .726, *η2p* = .000, 95% CI [0.00, 0.01], or an interaction between expecting a benefit and benefit type on perceptions of moral character, *F*( 2, 994) = .128, *p* = .880, *η2p* = .000, 95% CI [0.00, 0.00]. We found no support for the hypothesis that a) not expecting to receive emotional rewards are seen as more moral than those who expect to receive emotional rewards, and b) expecting emotional rewards has a divergent effect on perceptions of moral character compared to when expecting material and reputational rewards. In comparison, the target article found support for expected benefit (*F(*1, 280) = 34.7, *p* < .001, *η2p* = .11, 95% CI [0.05, 0.18]) and expected benefit × benefit type (*F*(1, 280) = 25.4, *p* < .001, *η2p* = .15, 95% CI [0.08, 0.23]) on perceptions of moral character. We provided a summary plot in Figure 7.

###### Figure 7 *Study 6: Moral character*

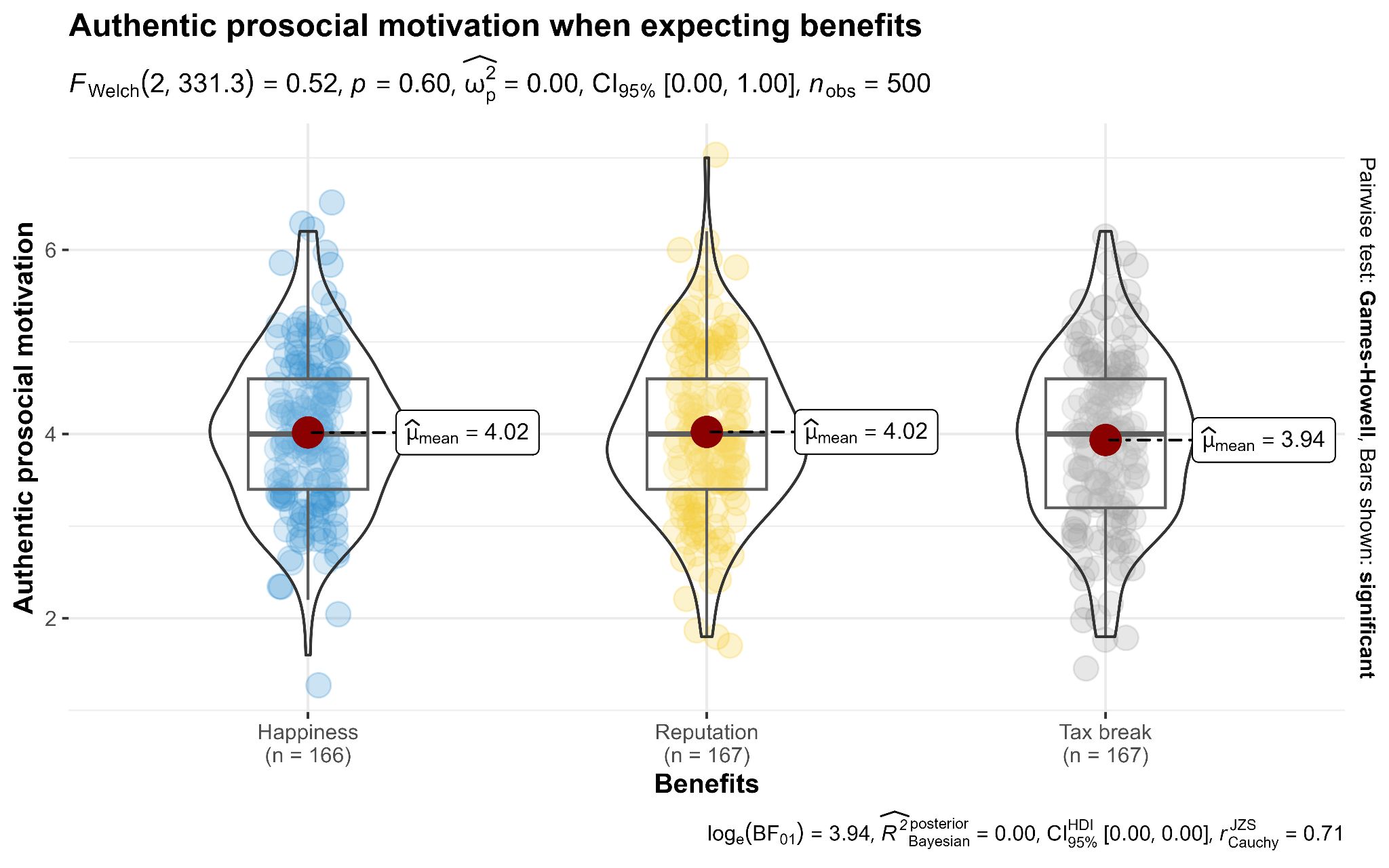
**

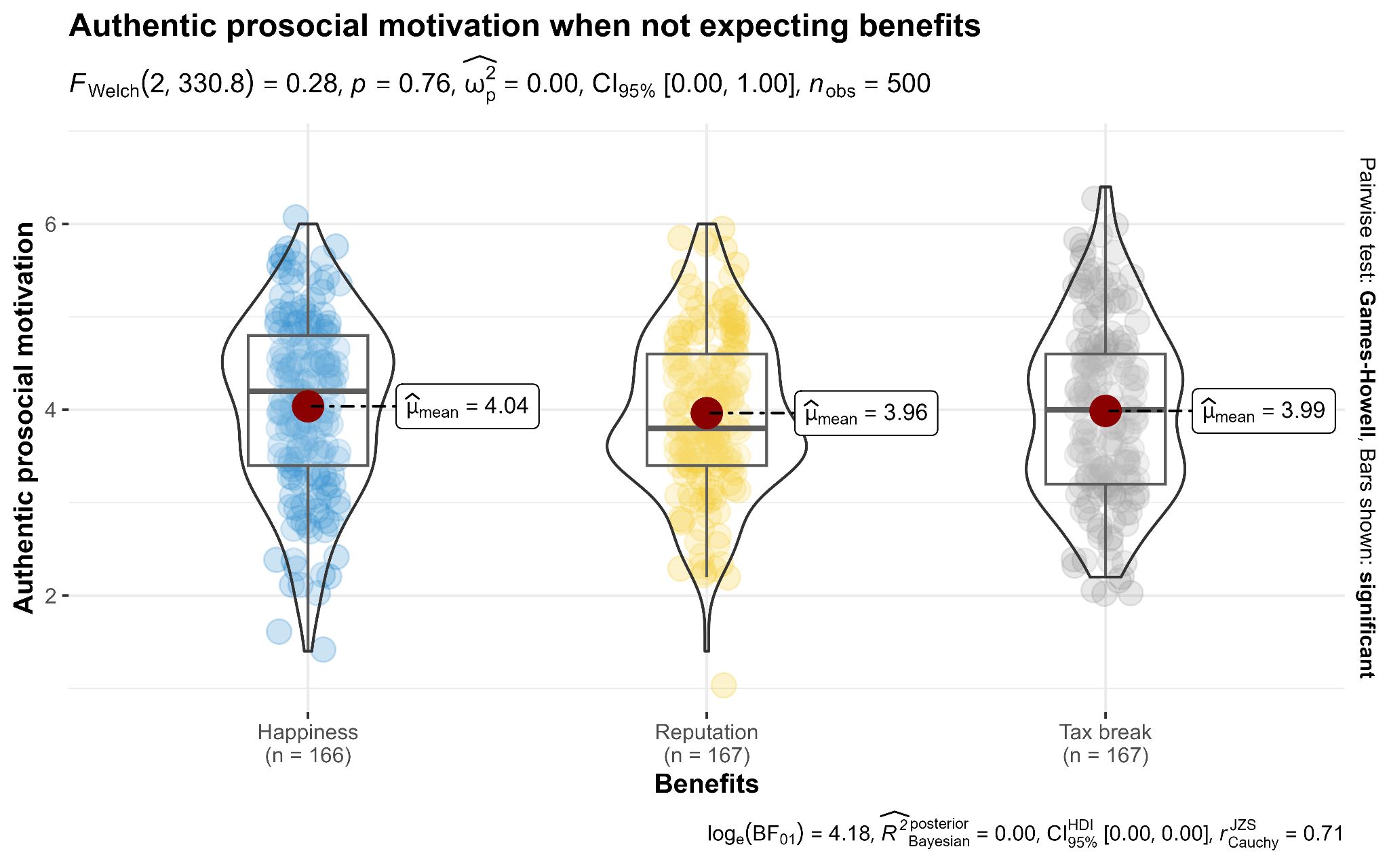
**

#### Authentic prosocial motivation

The two-way ANOVA found no support for an effect of expecting a benefit on judgements of authentic prosocial motivation, *F*(1, 994) = .310, *p* = .578, *η2p* = .000, 95% CI [0.00, 0.01], benefit type on judgements of authentic prosocial motivation, *F*(2, 994) = 2.63, *p* = .072, *η2p* = .005, 95% CI [0.00, 0.02], nor was there an interaction effect between expecting a benefit and benefit type on judgements of authentic prosocial motivation, *F*( 2, 994) = .069, *p* = .933, *η2p* = .000, 95% CI [0.00, 0.00]. We provided a summary plot in Figure 8.

###### Figure 8 *Study 6: Authentic prosocial motivation*

**

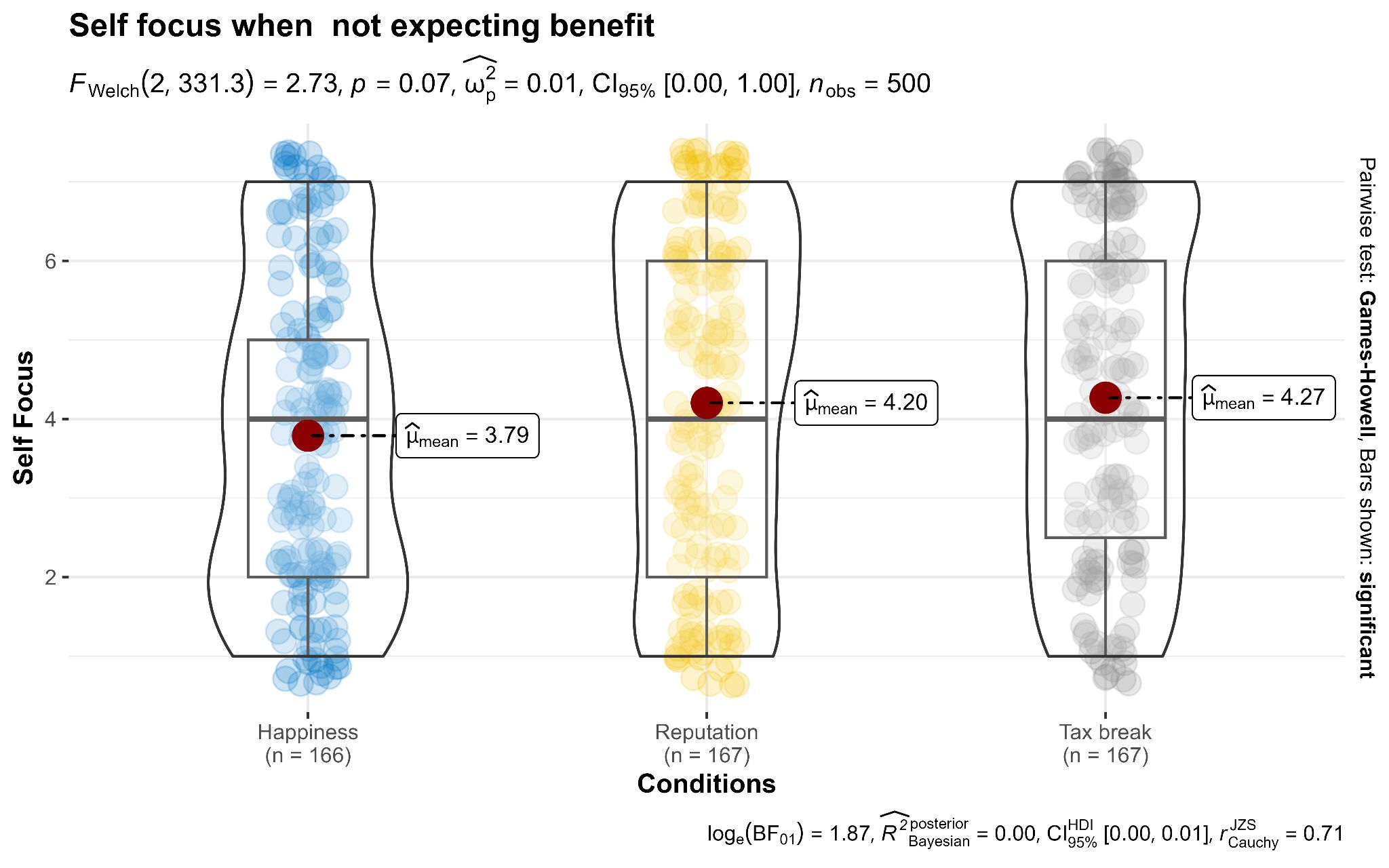
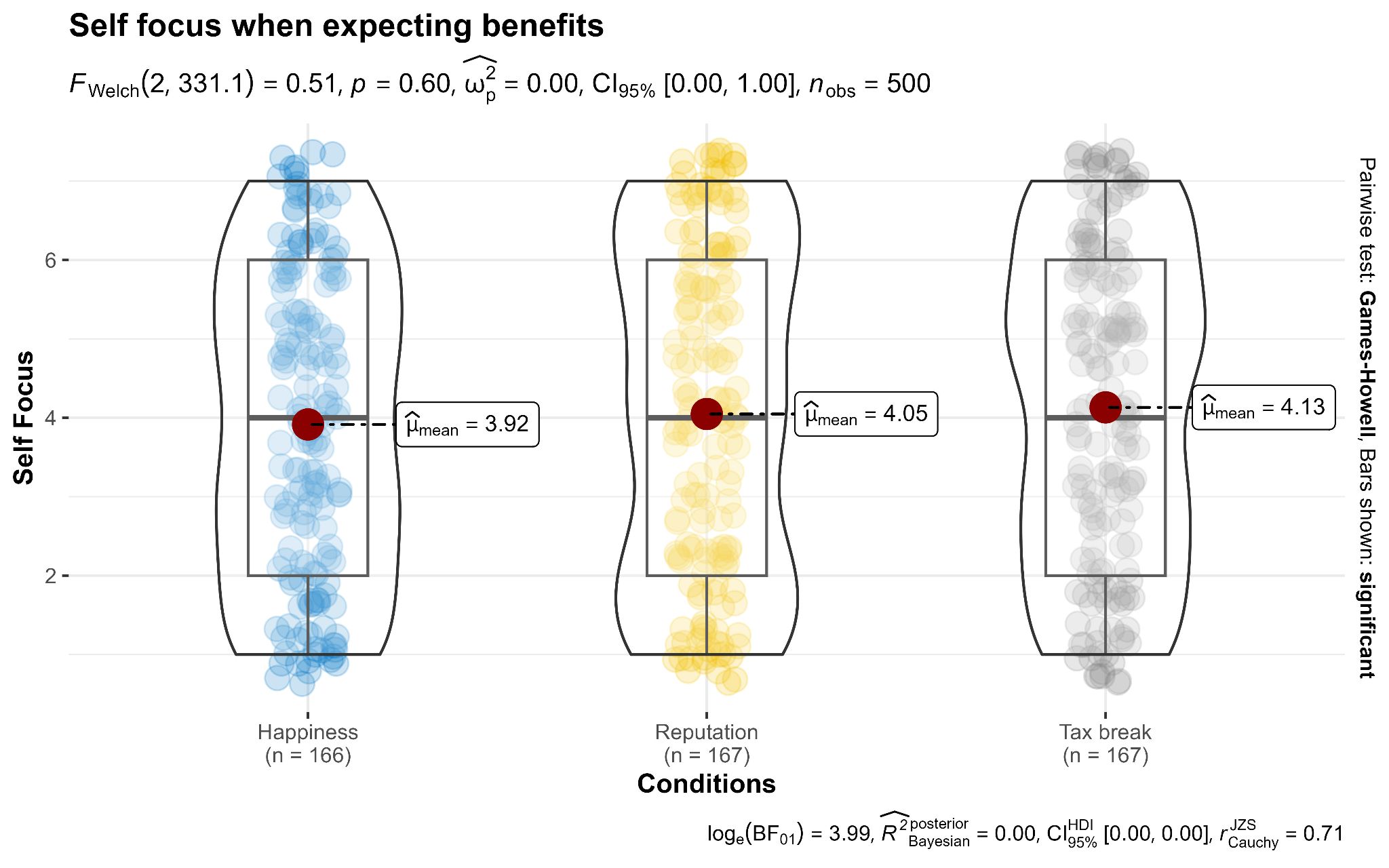
**

### Extension

#### Self-focus

We conducted a two-way ANOVA and did not find support for a main effect of expecting a benefit on perceived self-focus (*F*(1, 994) = 0.187, *p* = .666, *η2p* = .000, 95% CI [0.00, 0.01]), benefit type on perceived self-focus (*F*(2, 994) = 2.70, *p* = .068, *η2p* = .000, 95% CI [0.00, 0.02]), or an interaction (*F*(2, 994) = .502, *p* = .606, *η2p* = .001, 95% CI [0.00, 0.01]). We provided a summary plot in Figure 9.

###### Figure 9 *The effects of expected benefit and benefit type on perceived self-focus*

**

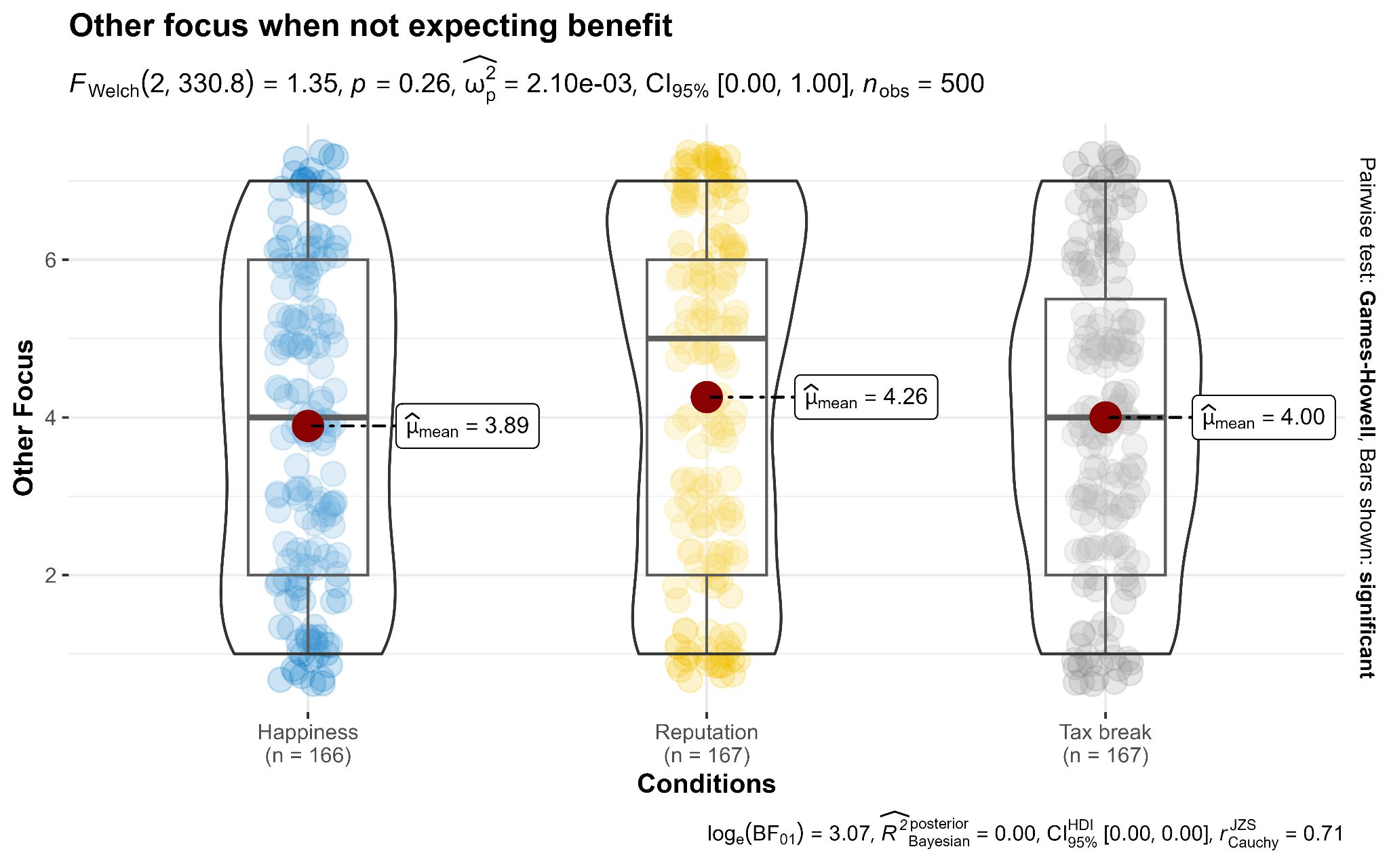
*Note:* Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

#### Other-focus

We found no support was for Hypothesis 5. We conducted as two-way ANOVA and did not find support for a main effect of expecting a benefit on perceived other-focus, *F*(1, 994) = .081, *p* = .775, *η2p* = .000, 95% CI [0.00, 0.00], benefit type on perceived other focus, *F*(2, 994) = .347, *p* = .707, *η2p* = .001, 95% CI [0.00, 0.01], nor an interaction effect between expecting a benefit and benefit type, *F*(2, 994) = .163, *p* = .850, *η2p* = .000, 95% CI [0.00, 0.00]. We provided a summary plot in Figure 10.

###### Figure 10 *The effects of expected benefit and benefit type on perceived other-focus*

**

**

*Note.* Created using ggbetweenstats in ggstatsplot R [Version 4.3.2] (Patil, 2021)

# Discussion

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

[Planned discussion for Stage 2: Possible weaknesses in the target’s design, and our adjustments and findings, such as regarding Study 6 conflation of outcomes and expectations and our adjustment to remove that and focus only on expectations.]

[Planned discussion for Stage 2: Based on comment by Dr./Prof. Romain Espinosa, we note that we will discuss issues of scale validity/reliability if scale reliabilities in the replication are below Cronbach’s alpha of .60. In such a case we also see value in a discussion of the challenges in how replications should approach issues of scale reliability thresholds.]

# Conclusion

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

# References

Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, Vol. 20, pp. 65–122). Academic Press. https://doi.org/10.1016/S0065-2601(08)60412-8.

Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 61*(3), 413–426. https://doi.org/10.1037/0022-3514.61.3.413..

Batson, C. D., & Shaw, L. L. (1991). Evidence for Altruism: Toward a Pluralism of Prosocial Motives. *Psychological Inquiry, 2*(2), 107–122. <https://doi.org/10.1207/s15327965pli0202_1>

Batson, C. D. (2010). Empathy-induced altruistic motivation. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature* (pp. 15–34). American Psychological Association. https://doi.org/10.1037/12061-001.

Ben-Shachar, M.S., Lüdecke, D., Makowski, D., (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software, 5*(56), 2815, [https://doi.org/10.21105/joss.0281](https://doi.org/10.21105/joss.02815)

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.

Buchanan, E., Gillenwaters, A.M., Scofield, J. and Valentine, K.D. (2022). *MOTE: Effect Size and Confidence Interval Calculator*. R package version 1.0.2. https://cran.r-project.org/web/packages/MOTE/MOTE.pdf.

Carlson, R. W., & Zaki, J. (2018). Good deeds gone bad: Lay theories of altruism and selfishness. *Journal of Experimental Social Psychology, 75*, 36–40. <https://doi.org/10.1016/j.jesp.2017.11.005>

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., 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.

Champely, S. (Developer), Ekstrom, C. (Developer), Dalgaard, P. (Developer), Gill, J. (Developer), Weibelzahl, S. (Developer), Anandkumar, A. (Developer), Ford, C. (Developer), Volcic, R. (Developer), & De Rosario, H. (Developer). (2017). *pwr: Basic functions for power analysis*. Software https://cran.r-project.org/web/packages/pwr/.

Cialdini, R. B. (1991). Altruism or Egoism? That Is (Still) the Question. *Psychological Inquiry, 2*(2), 124–126. https://doi.org/10.1207/s15327965pli0202\_3.

Cialdini, R. B., Darby, B. L., & Vincent, J. E. (1973). Transgression and altruism: A case for hedonism. *Journal of Experimental Social Psychology, 9*(6), 502–516. https://doi.org/10.1016/0022-1031(73)90031-0.

Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-based helping: Is it selflessly or selfishly motivated?. *Journal of personality and social psychology, 52*(4), 749. https://doi.org/10.1037/0022-3514.52.4.749.

Epley, N., Kumar, A., Dungan, J., & Echelbarger, M. (2023). A prosociality paradox: How miscalibrated social cognition creates a misplaced barrier to prosocial action. *Current Directions in Psychological Science*, *32*(1), 33-41. https://doi.org/10.1177/09637214221128016.

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

Hobbes, T., (1651). *Leviathan: Or the matter, form, and power of a commonwealth, ecclesiastical and civil*. London: A. Crooke.

Hoffman, M. L. (1981). Is altruism part of human nature? *Journal of Personality and Social Psychology, 40*(1), 121–137. https://doi.org/10.1037/0022-3514.40.1.121

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>.

Jordan, J. J., Hoffman, M., Nowak, M. A., & Rand, D. G. (2016). Uncalculating cooperation is used to signal trustworthiness. *Proceedings of the National Academy of Sciences, 113*(31), 8658-8663. https://doi.org/10.1073/pnas.1601280113.

Kelley, K., Kelley, M. K., & Imports, M. A. S. S. (2018). *The MBESS R Package. Computer software]. MBESS*. Retrieved from: https://CRAN. R-project. org/package= MBESS.

Klein, N., Grossmann, I., Uskul, A. K., Kraus, A. A., & Epley, N. (2015). It pays to be nice, but not really nice: Asymmetric reputations from prosociality across 7 countries. *Judgment and Decision Making*, *10*(4), 355-364. https://doi.org/10.1017/S1930297500005167.

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.

Levine, E. E., Barasch, A., Rand, D., Berman, J. Z., & Small, D. A. (2018). Signaling emotion and reason in cooperation. Journal of experimental psychology. *General, 147*(5), 702–719. https://doi.org/10.1037/xge0000399.

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.

Parrott, W. G. (2019). Emotions as signals of moral character. In U. Hess, & S. Hareli (Eds.) The Social Nature of Emotion Expression: *What Emotions Can Tell Us About the World* (pp. 161-177). Springer Cham. https://doi.org/10.1007/978-3-030-32968-6.

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.

Paxton, P., Velasco, K., & Ressler, R. W. (2020). Does use of emotion increase donations and volunteers for nonprofits?. *American Sociological Review, 85*(6), 1051-1083. https://doi.org/10.1177/0003122420960104.

Pfattheicher, S., Nielsen, Y. A., & Thielmann, I. (2022). Prosocial behavior and altruism: A review of concepts and definitions. *Current opinion in psychology, 44*, 124–129. https://doi.org/10.1016/j.copsyc.2021.08.021

Reed, L. I., Zeglen, K. N., & Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evolution and Human Behavior, 33*(3), 200-209. https://doi.org/10.1016/j.evolhumbehav.2011.09.003.

R Core Team (2023). 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).

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior, 31*(2), 87-94. https://doi.org/10.1016/j.evolhumbehav.2009.09.006.

Stearns, D. C., & Parrott, W. G. (2012). When feeling bad makes you look good: Guilt, shame, and person perception. *Cognition & emotion, 26*(3), 407-430. https://doi.org/10.1080/02699931.2012.675879.

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

VanBergen, N., Lurie, N. H., & Chen, Z. (2022). More Rational or More Emotional Than Others? Lay Beliefs About Decision‐Making Strategies. *Journal of Consumer Psychology, 32*(2), 274-292. https://doi.org/10.1002/jcpy.1244.

Van Kleef, G. A., De Dreu, C. K., & Manstead, A. S. (2004). The interpersonal effects of anger and happiness in negotiations. *Journal of personality and social psychology, 86*(1), 57. https://doi.org/10.1037/0022-3514.86.1.57.