PCIRR #373 Tsang 2006 rep-ext 2nd Revise and Resubmit: Response/reply to reviewers/decision letter

We would like to thank the editor and the reviewers for their useful suggestions and below we provide a detailed response as well as a tally of all the changes that were made in the manuscript. For an easier overview of all the changes made, we also provide a summary of changes.

Please note that the editor's and reviewers' comments are in bold with our reply underneath in normal script.

A track-changes comparison of the previous submission and the revised submission can be found on: https://draftable.com/compare/YZBODChGnkvL (or an export file on the OSF: https://osf.io/pqc5u)

A track-changes manuscript is provided with the file: "PCIRR-RNR2-Tsang2006-replication-extension-main-manuscript-trackchanges.docx" (https://osf.io/wvmsh)

Summary of changes

Below we provide a table with a summary of the main changes to the manuscript and our response to the editor and reviewers:

Section	Actions taken in the current manuscript
Introduction	Ed: We modified our description for the replication and extension analyses. We reframed the hypotheses and corrected the typographical errors.
	R2: We referenced Nelson et al. (2023).
Methods	Ed: We reiterated and clarified our decision to combine Studies 2 and 3 into one data collection in more sections. We improved on our description for the replication and extension analyses.
	R2: We clarified that we have sufficient power for moderation analyses under the section of "order effect". We stated that participants will see both Study 2 and Study 3 stimuli in our replication in Table 6.
Results	R2: We adjusted the titles for Figure 1 and Figure 2.
Reference list	Ed: We cited R and G*power and effect size guide.
	R3: We re-added references (e.g., Ames et al., 2004).
Supplementary materials	R2: We adjusted the comprehensive check under the section "Materials and scales used in the replication and extension experiment."

Note. Ed = Editor, R1/R2/R3 = Reviewer 1/2/3

We note that we are not familiar with the titles and ranks of the reviewers, and looking for that information proved tricky. To try and err on the side of caution, we refer to all reviewers with the rank Dr./Prof. We apologize for any possible misalignments and are happy to amend that in future correspondence.

Reply to Editor: Dr./Prof. Zhang Chen

> **I have now received comments from three reviewers who have also reviewed the original RR. We are overall all satisfied with the revisions that you have made, and think that the revised manuscript has improved much. Two reviewers have provided some further minor comments, and I too have noted some small issues while reading the manuscript again myself (see below). I wish you good luck with adressing these final comments, and I look forward to seeing the revised RR.**

Thank you for the reviews obtained, your feedback, and the invitation to revise and resubmit.

> **Minor comments:**

> **.1. Page 9: "In line with the difference in action tendencies between gratitude and indebtedness, they also found that participants were more likely to express willingness to return the favor if the benefactor did not communicate strong reciprocation expectations." You mentioned above that "gratitude leads people to thank their benefactor, whereas indebtedness leads people to try and return the favor", and that "higher expectations resulted in decreased gratitude yet increased indebtedness".**

> **Combining these two findings, I would expect that no strong reciprocation expectations -> increased gratitude and decreased indebtedness -> reduced willingness to return the favor, thus opposite to the claim in the first sentence?**

Thank you, this is important feedback. We accept and modified accordingly:

> Researchers have tried to distinguish the two emotions in different ways. In terms of their causes, Watkins et al. (2006) distinguished them by manipulating helpers' expectations of reciprocity, finding that higher expectations resulted in decreased gratitude yet increased indebtedness. In line with the difference in action tendencies between gratitude and indebtedness, they also found that participants were more likely to express willingness to return the favor if the benefactor communicated weaker reciprocation expectations (reciprocation negatively associated with reciprocation expectations). However, this predicted outcome is arguably counter to combining Tsang's (2006) and Algoe et al.'s (2010) theoretical paradigms, which together suggest that beneficiaries would be less likely to reciprocate if benefactors held weaker reciprocation expectations (reciprocation positively associated with reciprocation expectations).

Table 1: Hypotheses 4, 5 and 4+5 are said to be re-analyses of hypotheses 2, 3 and 2+3, but they do not correspond to each other (4 and 5 are the same).

Thank you for catching that. That was an oversight.

Hypothesis 5 was previously written as:

[Regression complementary analysis] Benevolent favors result in more gratitude than selfish favors, even after controlling the magnitude of favor.

Yet, ofcourse, it was meant to be about indebtedness, and so we corrected it to :

[Regression complementary analysis] Benevolent favors result in more indebtedness than selfish favors, even after controlling the magnitude of favor.
[Reframed from the target article's null hypothesis]

Table 1: Hypothesis 7c: "Gratitude is higher in Ambiguous condition compared to Selfish condition", but the effect size Cohen's d is negative.

Yes, this is a good point, thank you. We previously simply followed directly from their reported effects:

Participants felt significantly more grateful in the Benevolent Motives condition compared to the Ambiguous Motives condition [$t(41.63) = 2.07, p < .05$], <u>and marginally less grateful in the Ulterior Motives condition compared to the Ambiguous Motives condition [$t(56) = -1.84, p = .07$]</u>

Which we calculated to convert to effect size as: $d = -0.49$ [-1.01, 0.04].

Given your feedback, it does make more sense that the effect size would be in the same direction as the hypotheses, and so we changed it to the following: $d = 0.49$ [-0.04, 1.01].

Table 1: Is there a reason why hypothesis 7c is about ambiguous versus selfish, whereas hypothesis 8c is about benevolent versus selfish? Later on, 7c and 8c are compared, but the two do not really seem comparable.

Thank you for catching that, this was an oversight. Hypothesis 8c is about "Ambiguous condition compared to Selfish condition". This has been adjusted.

Page 19: Please also cite R and G*Power to give credits to their developers.

We added the relevant citations and references to our manuscript under the section "Effect size and confidence intervals":

Effect size and confidence intervals were calculated with R (Version: 4.1.2; R Core Team, 2020) with the help of a guide by Jané et al. (2024), and power analyses were then conducted with a combination of R and GPower (Version 3.1.9.6; Faul et al., 2007) for the factors that the authors found support for in the target article (flagged as significant results).

> **Page 29: "There were 16 effect sizes calculated from the target study (see Table 2)". Should be Table 1?**

Thank you, we adjusted it to point to Table 1. Much appreciated.

> **There is some inconsistency in how the 'selfish' condition is named throughout the manuscript (sometimes called ulterior). It's better to use a consistent label.**

Thank you for bringing this up. This is an issue with the target article, and we previously discussed how to best address the issue amongst ourselves, considering the ups and downs of standardizing. We finally decided against it, to better mirror the target article.

To explain the issue, in the original article, the "selfish condition" was the label used in their Study 2, whereas the "ulterior condition" was the label used in their Study 3. We were worried that choosing either the word "selfish" from Study 2 or "ulterior" from Study 3 may confuse those comparing the replication to the target article. Therefore, adhering to the wording adopted in the original article may be the best approach to accurately present the original experimental design as a replication study.

Given your feedback, we can see how that might be confusing. We therefore revisited this issue, and realized the inconsistencies and possible confusion, and so we decided to change "selfish" and "ulterior" to "selfish-ultieror" throughout the manuscript for the condition label. We hope that is clearer and a satisfactory solution.

> **Table 6: It may be good to explicitly mention (again, perhaps under Procedural details) that the target article conducted Studies 2 and 3 separately, whereas here the two studies are combined into one data collection batch.**

We modified the paragraph and explicitly mentioned in the procedure section that the studies were combined into one data collection batch:

> "Following consent and qualification questions, participants completed the two studies as replications of Studies 2 and 3 in the target article, in random order. In contrast to the

original paper where the two studies were conducted separately, we combined the two studies into a single data collection."

> **Page 32, Replication: Extension analyses: I can understand the analyses, but I find the description a bit difficult to follow. It may help the readers if you could explicitly specify which ANOVAs were conducted, something like "mixed ANOVA with intention condition (benevolent versus selfish in Study 2, and benevolent versus selfish versus ambiguous in Study 3) as a between-subject factor, and emotion type (gratitude versus indebtedness) as a within-subject factor, and the reported emotion strength as the dependent variable".**

Thank you for the suggestion. We modified our description for the analyses based on your recommendation:

> "In both studies, the comparison between gratitude and indebtedness was done by comparing signals, in which one intent affected gratitude but no indebtedness, which we reframed to a comparison of the effects of the two dependent variables. To complement that analyses, we will conduct extension analyses of a 2-way mixed ANOVA, with intention conditions as a between-subject factor (benevolent versus selfish in Study 2, and benevolent versus selfish versus ambiguous in Study 3), emotion type as a within-subject factor (gratitude versus indebtedness), and emotion ratings as the dependent variable."

> **Page 33: Is there a rationale for why the alpha level is adjusted to .005 (in Order effects) and .001 (in Outliers and exclusions)?**

Thank, that does require clarification. We therefore added the following:

> We plan this as a second order analysis only after the "Order effects" analyses above also fail to find support for the analyses.

Therefore, if we run the initial order effect analyses (if replication fails), we set alpha to .005. Then if we again fail to find support, we will then follow with outliers and exclusions, and will therefore aim for an even lower alpha of .001.

Reply to Reviewer #1: Dr./Prof. Sarahanne Miranda Field

> **I have read both the authors' response to reviews and the revised manuscript. I am pleased to see that much work has been done to improve the proposed study protocol, and I am happy with how the authors responded to my comments. I am satisfied that a study based on this protocol would be of sound methodological quality, and would provide enough detail to be reproducible by other parties. I look forward to seeing the study when it is complete.**

We sincerely appreciate your positive feedback and support.

Reply to Reviewer #2: Dr./Prof. Jo-Ann Tsang

> **I appreciate the authors' responses to the reviews, they helped clarify some confusion and concerns I had with the research. Although I am hesitant about the within-subjects nature of the registered report study, as long as the authors collect enough data to examine order effects (which it appears they will), then I have no problem with it. Below are some small suggestions to further improve the research.**

Thank you very much for your feedback. We will ensure that that order effects would not hamper the reliability and credibility of the results and conclusions we made.

> **.1.**
> **- on p. 17 of the marked document, item #9, it might be clearer if, rather than using the label "selfish intentions" the authors used "benevolent motivations" so that it reads "Ratings of benevolent motivations are associated with gratitude . . ." This suggestion also applies to item 10, and item 9 + 10.**

Thank you, that is a good point.

The scale in Tsang (2006) for the helper intention manipulation check went from benevolent (1) to selfish (7) so we framed it as selfish intentions. Yet, we agree that this would be better described as helper intentions/motivations, and made the adjustment.

> **.2.**
> **- related to the authors' proposed extension looking at indebtedness and gratitude on reciprocity, Nelson et al 2023 also looked at gratitude, indebtedness, and prosociality (https://doi.org/10.1080/17439760.2023.2190926)**

Thank you for suggesting this article. We now reference this paper when discussing our extension under the section of "Exploratory extension":

> To the best of our knowledge, there has been no research examining the impact of helper intention on reciprocation magnitude. A related research by Nelson et al. (2023) examined the influence of gratitude and indebtedness on prosocial behaviors, and our extension would complement that by examining helper intentions and using a different paradigm and measures.

.3.
- around p. 20, the authors report sensitivity and power analyses. The authors may want to explicitly state that the number of participants they plan on recruiting are more than enough to provide enough power for moderation analyses looking at order effects on the presentation of Study 2 vs. Study 3.

We appreciate the suggestion. We added a note under the section of "Order effects" to point out we have enough power for the moderation analyses:

> Our planned sample size (1000) is large enough to provide sufficient statistical power to conduct moderation analyses and examine the order effects, if needed.

.4.
- on p. 26, on the item "Study 2 and 3 are conducted seperately", many of the points raised in the "Reasons for change" column are unrelated to reasons for changing the original study to within-subjects. For example, the reasons "To reduce the order effect" and "to avoid the influence of decline to particular studies" are not the reasons why the authors moved to a within-subject design (although they are reasons for counterbalancing). The only reason in that section that explains a change from the original study is the statement, "to find potential consistency within participants' answers (whether an answer is predictive of another answer". The other reasons, while good reasons for counterbalancing, are not relevant to the change from the between Ss nature of the original studies, and the within-Ss nature of the proposed study.

We appreciate the feedback, and it has helped us realize that explaining the reasons for change for two items together can be confusing. To improve clarity, we separated the two items into distinct rows. Below is the updated table for your reference:

Table 5

Comparison of target article versus replication

	Target article	**Replication**	**Reasons for change**
Study design	Participants completed the studies with pen and paper in the laboratories.	Participants completed the studies with an online survey.	To reach more and a wider variety of participants; to conduct the studies with lower cost and higher efficiency.
Sample characteristics	Sample size: Study 2: 92; Study 3: 86	$N = 1000$ (TBC)	Generalizability of results by including a more and larger variety of participants.
	Geographic origin: Undergraduates studying at Baylor University	Students at online research platform *Prolific*	
Procedure	Items on gratitude and indebtedness were not randomized	Items on gratitude and indebtedness were randomized	To reduce the order effect.
	Studies 2 and 3 were conducted separately	Studies 2 and 3 were conducted with the same participants in one setting	Potentially explore consistency in participants' answers across studies (whether an answer in one study is predictive of an answer in the other study) and order/decline effects.
		The order of Studies 2 and 3 is randomized for the participants.	To explore and address potential order effects. Allows subsequent separate analysis on participants who took Study 2 / 3 as their first presented study.
Procedure	No comprehension check	We added comprehension checks for Studies 2 and 3	To ensure that the participants read and understood the instructions and scenarios.
Conditions	No Change	No Change	NA

.5.

- on p. 37 on the table for the classification of replication, would it be appropriate to include in procedural details the fact that participants will see both Study 2 and Study 3 stimuli in the new research?

Thank you for raising this point. We added extra details under the design facet in the table and mentioned that the same participant would answer both Study 2 and Study 3 in a single data collection:

Table 6
Classification of the replication, based on LeBel et al. (2018)

Design facet	Replication	Details of deviation
Effect/hypothesis	Same	
IV construct	Same	
DV construct	Same	
IV operationalization	Same	
DV operationalization	Same	
Population (e.g., age)	Similar	<u>Target article:</u> The study recruited students from Baylor University in the United States <u>Current replication:</u> The study targeted students on the online research platform Prolific
IV stimuli	Similar	<u>Target article:</u> Two groups of subjects were recruited to receive stimuli from Study 2 and Study 3 respectively <u>Current replication:</u> The same participants answered both Study 2 and Study 3
DV stimuli	Same	
Procedural details	Similar	1. <u>Target article:</u> One comprehension check was used <u>Replication:</u> One extra comprehension check was added 2. <u>Target article:</u> Items on action tendency, Positive and Negative Affect Schedule (PANAS), Gratitude Resentment and Appreciation Scale (GRAT) and The Gratitude Questionnaire – Six Item Form (GQ-6) questionnaire were not randomized. <u>Replication:</u> These items on the questionnaire were randomized
Physical settings	Different	<u>Target article:</u> Participants completed the studies with pen and paper in a laboratory setting <u>Replication:</u> Participants completed the studies online through Prolific.
Contextual variables	Different	Different time and context.
Replication classification	Close replication	

.6.
- on p. 46 the title for Figure 1 needs to be fixed.
- on p. 47 the title for Figure 2 needs to be fixed.

Thank you for catching that. We corrected the titles:

Figure 1

Study 2 benevolent condition: Association between gratitude and indebtedness
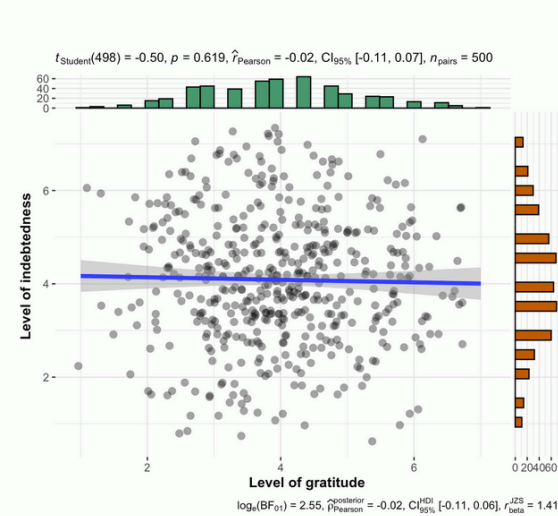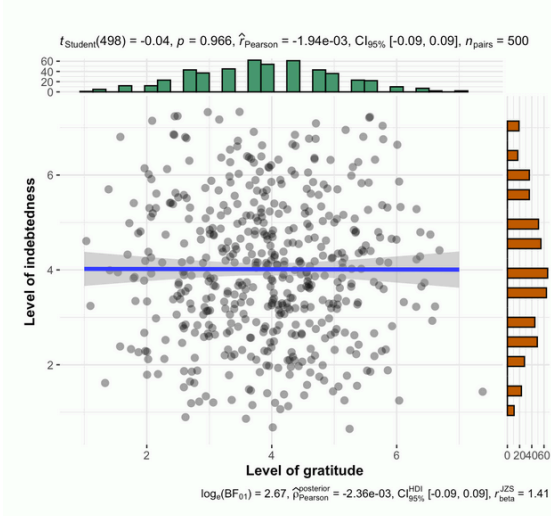
$t_{Student}(498) = -0.50$, $p = 0.619$, $\hat{r}_{Pearson} = -0.02$, $CI_{95\%}$ [-0.11, 0.07], $n_{pairs} = 500$

$\log_e(BF_{01}) = 2.55$, $\hat{\rho}_{Pearson}^{posterior} = -0.02$, $CI_{95\%}^{HDI}$ [-0.11, 0.06], $r_{beta}^{JZS} = 1.41$

Figure 2

Study 2 selfish-ulterior condition: Association between gratitude and indebtedness

$t_{Student}(498) = -0.04$, $p = 0.966$, $\hat{r}_{Pearson} = -1.94e-03$, $CI_{95\%}$ [-0.09, 0.09], $n_{pairs} = 500$

$\log_e(BF_{01}) = 2.67$, $\hat{\rho}_{Pearson}^{posterior} = -2.36e-03$, $CI_{95\%}^{HDI}$ [-0.09, 0.09], $r_{beta}^{JZS} = 1.41$

.7.
- In the authors' response letter, the third comprehension check for Study 3 appears to have incorrect answers. "a" (I know without doubt it is because my friend wanted to borrow my car.) should be selfish, "b" (It is not clear about the two being related, but the weekend after helping me this friend asked to borrow my car) sho- on uld be ambiguous, and "c" (My friend is really concerned about me) should be benevolent.

Thank you for catching that. These were implemented correctly in the Qualtrics but were not updated in the supplementary.

We reviewed the comprehensive checks in Qualtrics once again and have confirmed that they were correct. The updated information in our supplementary document is as follows:

Study 3

We added three comprehension-check questions for Study 3. One was directly extracted from the original study, with additional two new questions. Participants were not allowed to proceed to the next page unless they answered the questions correctly. The answers were presented in randomized order.

1. How much money did the friend offer to give to help pay for the textbooks?

    Ans: 200

2. What was the favor offered in the scenario?
    a. A birthday gift
    b. Helping you with your homework
    c. Paying for the textbooks for you

    Ans: C. Paying for the textbooks for you

3. According to the text: Why is your friend offering to help you?
    a. I know without doubt it is because my friend wanted to borrow my car
    b. It is not clear about the two being related, but the weekend after helping me this friend asked to borrow my car
    c. My friend is really concerned about me

    Ans: a. for *Selfish Condition*

    b for *Ambiguous condition*

    c for *Benevolent Condition*

I look forward to seeing the authors' completed research.

Thank you for your research, and for your reviewing help and support.

Reply to Reviewer #3: Dr./Prof. Cong Peng

> **Thank the authors very much for clearly addressing my concerns in detail, which is impressive. I think the current manuscript is much more improved.**

Thank you. We very much appreciate the detailed comments and suggestions.

> **.1.**
> **I still have one small concern about the extension with Watkins et al (2006). I wonder whether perceived expectation for reciprocity is also a manipulation check rather than a DV. In Study 3 in the Selfish condition, the intention seems out of expectation for reciprocity. And in the vignette it actually has explicitly mentioned the higher expectation for reciprocity: "in order to borrow your car next weekend".**

Thank you for raising this. Yes, this is an interesting point, the manipulation includes both an element of manipulating intentions and manipulating reciprocity.

We treated "perceived expectations for reciprocity" as an extension dependent variable given that the target article did not intend the manipulation as that of reciprocity, and although it did measure helper intentions as a manipulation check it did not include this measure. We added "perceived expectations for reciprocity" in our replication study to test directly how these conditions impact reciprocity, to allow for a more comprehensive examination of the influence of gratitude and indebtedness on reciprocity tendencies, and also examine associations between the dependent variables and checks.

We added to Extensions->Perceived expectations for reciprocity a note addressing this issue:

> We note that although the target article did not set off to manipulate expectations, Study 3 did vary expectations for reciprocity between the conditions with specific mention of such expectations in the selfish condition. As one reviewer noted, this measure could be considered a manipulation check examining the impact of making that expectation explicit in one of the conditions.

We also added a planned discussion of this point:

> [Planned discussion in Stage 2 following Dr./Prof. Cong Peng comment: We will discuss the expectations of reciprocity extension and its role as a possible manipulation check, given that the scenario also explicitly mentioned expectations in the selfish condition.]

.2.

Besides, some minor points. I saw a list of α in the Measures. What does α mean here? If they are Cronbach's α, they seem incorrect (all below 0.1). If they are not, please specify.

The α in the Measures section represents Cronbach's α values. All values were calculated from the simulated random data generated by Qualtrics for demonstration purposes. Given that it is basically random noise, values have to be small, if there were bigger that would be an indication of a problem.

These values will be updated with the results of the real data once we collect the actual data.

.3.

And Ames et al. (2004) was listed in the text (p.10) but not in the reference lists.

Thanks for catching that. We added the relevant reference:

Ames, D. R., Flynn, F. J., & Weber, E. U. (2004). It's the thought that counts: On perceiving how helpers decide to lend a hand. *Personality and Social Psychology Bulletin*, 30(4), 461-474. https://doi.org/10.1177/0146167203261890

Good luck with your research!

Thank you for your feedback and support.