Received: 30 November 2022 | Accepted: 8 February 2023

DOI: 10.1002/pchj.641

ORIGINAL ARTICLE

PsyCh Journal

Social incentives foster cooperation through guilt aversion: An effect that diminishes with primary psychopathic traits

Xiaoyan Wu[1] | Ruida Zhu[2] | Xu Gong[3,4] | Yuejia Luo[1,3,4] | Chao Liu[1]

[1]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

[2]Business School, Beijing Normal University, Beijing, China

[3]Center of Brain Disorder and Cognitive Sciences, College of Psychology, Shenzhen University, Shenzhen, China

[4]Shenzhen Key Laboratory of Affective and Social Cognitive Science, Shenzhen University, Shenzhen, China

Correspondence
Chao Liu, Center of Brain Imaging, Beijing Normal University, No.19 Xinjiekouwai Street, Haidian District, Beijing 100875, China.
Email: liuchao@bnu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 32271092 and 32130045; Major Project of National Social Science Foundation, Grant/Award Number: 19ZDA363

Abstract

The association between primary psychopathic traits and non-cooperative behaviors is well-identified. There is a lack of studies on how to motivate cooperative behaviors in individuals with primary psychopathic traits. This study investigated the effects of monetary incentives and social incentives on promoting cooperation in healthy adults with varying primary psychopathic traits. Participants played a one-shot public goods game (PGG) with other anonymous players in three different contexts: a social incentives context where participants' decisions would be judged by others, a monetary incentives context where participants' decisions would result in winning or losing money depending on their contributions, and a control condition where no additional incentives were implemented. We found that, compared to the control condition, both monetary and social incentives significantly improved participants' contributions to the public project—an indicator of cooperative behavior. However, the association between higher primary psychopathic traits and less cooperation was only observed in the context of social incentives. Computational modeling further revealed that this effect can be explained by the diminishing guilt aversion when participants deliberately violated their inferred expectations of themselves from others' perspectives. This study found that social incentives can encourage cooperative behaviors in non-clinical psychopathy, and identified the mental processes navigating this effect.

KEYWORDS
cooperation, guilt aversion, monetary incentives, primary psychopathic traits, social incentives

INTRODUCTION

Psychopathy is a severe personality disorder involving deficits in social–emotional processing. The main characteristics of psychopathy include egocentricity, cheating, and lack of empathy, affection, remorse, and guilt (Hare, 1998). The connection between psychopathic personality and low cooperative behaviors has been widely reported in both clinically diagnosed psychopaths and healthy individuals (Chang et al., 2011; Johnston et al., 2014; ten Brinke et al., 2015). Clinically diagnosed psychopaths have difficulty sustaining long-term reciprocal relationships (Hare, 1998). They are more likely to show competitive, uncooperative behaviors in social interactions (Mokros et al., 2008). There are two subtypes of psychopaths that have been developed and widely identified: the "primary" and "secondary" psychopathic traits (Skeem et al., 2007), and studies have suggested these two kinds of psychopathic traits have different associations with cooperative behaviors. A study of clinical patients found that the behaviors of secondary psychopaths were more like those of healthy controls after cooperation and communication, whereas primary psychopaths exhibited less cooperation under the same condition (Widom, 1976). In the non-clinical sample, individuals with higher total and primary psychopathy scores were significantly correlated with fewer cooperative behaviors in the economic game. However, the association between secondary psychopathy scores and cooperative behaviors was not significant (Rilling et al., 2007). These findings underscore a specific role of primary psychopathic traits on non-cooperative behaviors. Some studies suggested the distinct personality structures (Cima & Raine, 2009) and evolutionary explanations (Glenn et al., 2011) for these two subtypes of psychopaths. Primary psychopathy is typically regarded as arising from genetic predisposition, which may include some inherent deficits, such as decreased emotional processing ability and reduced capacity for empathy (Lee & Ashton, 2005; Mealey, 1995); whereas secondary psychopathy is thought to have environmental or psychosocial causes, like

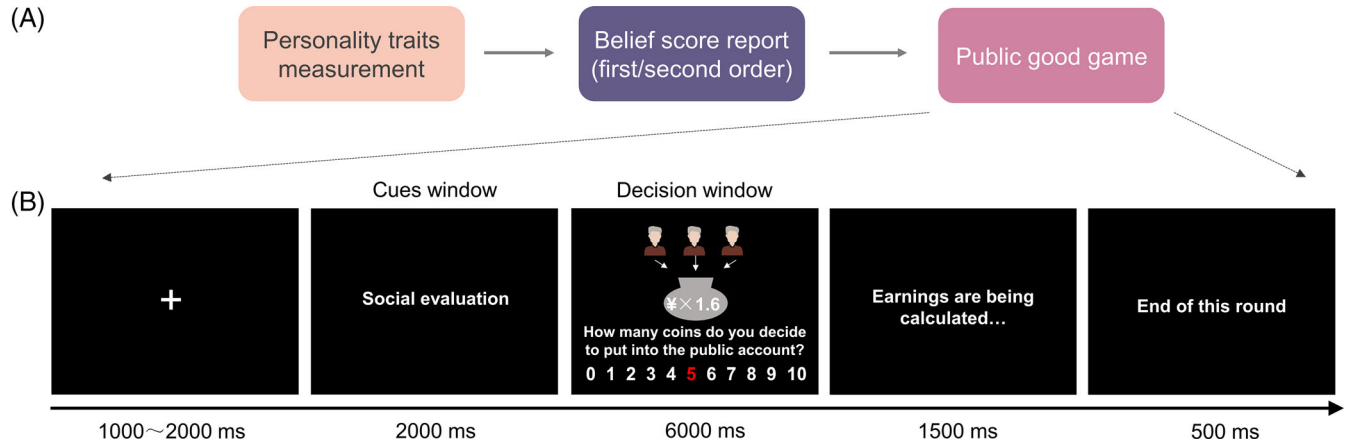

FIGURE 1 Experimental procedure and one trial illustration. (A) Research schema of the present study. All participants completed the experiment in a quiet behavioral laboratory. (B) One trial illustration. Each trial was started with a jitter (1000–2000 milliseconds), after which the cue word was presented (2000 milliseconds) to indicate the context of the current round. The cue words "social evaluation," "monetary incentives," and "impunity" were presented to indicate the contexts of social incentives, monetary incentives, and control, respectively. Thereafter, a decision window was shown during which participants were asked to make their decision within 6000 ms, followed by an ending window (500 ms) to indicate the end of the current trial. Notably, in the decision window, the red number was presented randomly across trials at the beginning to avoid any default hint. Participants were asked to press the left (right) button to move the red number to the left (right) option and press the enter button to submit the selection when ready.

exposure to parental abuse and neglect, substance abuse, and childhood trauma (Koenigs et al., 2010).

Although a relationship between non-cooperation and primary psychopathic traits has been frequently reported, few studies have explored ways to motivate these people to be more cooperative. There is some evidence that primary psychopaths' cooperative behaviors can be driven by the fear of monetary loss. Koenigs et al. (2010) found that, compared to secondary psychopaths and controls, primary psychopaths made lower offers in the Dictator Game (DG). However, in the Ultimatum Game (UG), where the receiver has the power to reject the offer and therefore participants face a threat of monetary loss, the offers made by psychopaths and controls were no different. Consistently, in repeated interactions where each player's payoff is dependent on mutual cooperation, the cooperative behaviors were not related to the psychopathic traits (Rilling et al., 2007). However, this relationship does not hold in one-off situations, where further reciprocity is not possible, and primary psychopathic traits are associated with fewer cooperative behaviors (Curry et al., 2011). Encouraging cooperative behaviors in people with high psychopathic traits in one-off situations is of great importance, as we are regularly exposed to strangers and have to deal with one-off interactions in our daily life. Monetary consequences may incentivize cooperative behavior in psychopaths, but more research is needed to determine the direct effects. Additionally, it is unclear how social consequences may impact cooperative behavior in this population. It is important to investigate the effect of the social consequences on cooperative behaviors, as in human daily life, behaviors are often linked to social consequences (such as the appreciation or dislike of others) instead of direct monetary gain or loss (Wu et al., 2016).

The purpose of this study was to investigate the effects of social incentives (others' evaluation) and of monetary incentives on cooperative behaviors in adults with different levels of psychopathic traits. Given that antisocial tendencies exhibited by non-clinical psychopaths have been deemed a potential threat to social stability (Anderson, 1999; Levenson et al., 1995), we selected healthy adults as our research sample. The psychopathic personality trait was assessed by the Levenson Self-report Psychopathy Scale (LSRP; Levenson et al., 1995), which has been widely used to assess psychopathic personality in a nonclinical population. In this study, participants played a modified version of the public goods game (PGG) with three anonymous players in each round (one-shot game). The game was played in three different contexts: a social incentives context, a monetary incentives context, and a control context. In the social incentives context, participants were evaluated by others based on their decisions; in the monetary incentives context, participants had the potential to gain or lose money depending on their contributions; and in the control context, no additional incentives were employed (see Figure 1). Participants' contributions in each context were regarded as an indicator of cooperative behaviors. We expected that both monetary and social incentives, compared with the control, could increase cooperation and sought to test their associations with psychopathic traits.

Previous studies suggested that the social dysfunctions of psychopathic personality may be due to these individuals' deficits in mentalizing, and thus they have difficulty following social expectations (Blair, 2007; Harenski et al., 2010; Koenigs et al., 2012). Based on that, we derived the following hypothesis:

> **Hypothesis 1.** Participants with high psychopathic traits may distort expectations about other players' decisions in PGG.

Participants' ability to build beliefs about others' expectations was assessed via self-reported scores. The self-reported

scores were assessed by asking participants to subjectively report their expectations regarding the decisions of others and others' expectations, which draws on previous studies (Chang et al., 2011; Gong et al., 2019). Conversely, a recent study found that individuals with higher psychopathic traits were correlated with reduced repayment in the trust game, but not with their self-reported scores about others' expectations (Gong et al., 2019). The findings suggested that individuals with higher psychopathic traits were able to understand social expectations, but that they did not follow them to make decisions that met the expectations of others. This result motivates our second hypothesis:

> **Hypothesis 2.** The social dysfunction of participants with high primary psychopathic traits is not due to mentalizing deficits but to their reluctance to follow social expectations, and therefore we cannot observe the association between psychopathic traits and self-reported scores.

We constructed models with different assumptions about how participants calculate the subjective utilities of cooperation and make decisions. Model 1 assumed that participants make decisions by trading off the interests of themselves (self-interests) and the interests of all four players in that round (collective-interests); we called this model the self-collective interests integration model. This model was based on the study of Park et al. (2011) and has been adapted for use in cooperative decision-making (Park et al., 2019). Under the assumption of the self-collective interests integration model, those who prioritize collective benefits over individual benefits will be more likely to cooperate and contribute more in PGG, and vice versa. Model 2, the guilt aversion model, assumed that people are motivated by self-interest and a desire to avoid the guilt that comes from disappointing others. This guilt can be relieved by making choices that align with others' expectations, such as increasing cooperation. The guilt aversion model was adopted from previous studies (Chang et al., 2011; Gong et al., 2019). A guilt aversion estimated from modeling reflects the degree to which participants are averse to the anticipated guilt of disappointing other players. The higher value of guilt aversion indicates that participants were more likely to conform to others' expectations and avoid guilt. If Hypothesis 2 held, that is, if the social dysfunction of participants with high primary psychopathic traits was due to their reluctance to follow social expectations, we expected to observe that participants with high psychopathic traits would exhibit less guilt aversion.

METHODS

Participants

A total of 60 adults (mean age = 21.68 ± 2.35 years; 19 males) participated in the present study. The sample size was determined by G*Power 3.1.9.7 (Faul et al., 2007) before the experiment. We set the probability of type I error (0.05), expected effect size ($f = 0.20$), power ($1-\beta = 0.90$), and correlation among repeated measures (0.5) which determined the minimum sample size to be 56 participants. The correlation among repeated measures is calculated by the correlations between predictors and dependent and the matrix of correlations among the predictors (Faul et al., 2007). We set the value of the correlation among repeated measures to be 0.5, which is based on previous work (Verma & Verma, 2020). We recruited slightly more participants to avoid insufficiency in sample size due to the possibility of data collection failure (e.g., technical problems or participant apathy or disinterest). The sample size of 60 participants in the current study would be sensitive to the minimum effects of Cohen's $f = 0.196$ with 90% power (alpha = 0.05, correlation among repeat measures = 0.5). Prior to the experiment, participants were assessed to ensure that they had no history of psychiatric disorders or substance abuse, had not suffered from any major medical illness in the last 6 months, and had normal or corrected-to-normal vision in a self-reported manner. Those who had majored in psychology or economics were excluded from participating. Informed consent was obtained from all individual adult participants included in the study.

Materials

The psychopathic personality trait was assessed by the LSRP (Levenson et al., 1995), which has been widely used to assess psychopathic personality in a nonclinical population. The Cronbach's alpha is a measure of internal consistency that is used to assess the reliability of tests and measures. In LSRP, the Cronbach's alpha achieved a score of .82 (Falkenbach et al., 2007). We used a Chinese version of the LSRP (Zheng, 2011). The LSRP contains 26 items and has a similar four-point Likert scale format. This measure is divided into factor analytically derived primary and secondary psychopathy subscores, which also roughly parallel Factor 1 and Factor 2 of the Psychopathy Checklist-Revised (PCL-R; Hare, 2003), respectively. The primary psychopathy items assess a selfish, uncaring, and manipulative posture toward others, and the secondary psychopathy items assess impulsivity and a self-defeating lifestyle. The measurements of LSRP scores for each subject are reported in Table S5.

The public goods game

The current study used an adapted version of the PGG (Figure 1). Participants played PGG with other anonymous players in three different contexts. In the context of the monetary incentives, participants were instructed that their performance would determine whether they received a financial reward or punishment. Specifically, if a participant invested fewer tokens than the average contribution of the other three players in a round, they would be penalized two tokens for each token they contributed to the public pool. Conversely, if they invested more than the average, they would be rewarded

two tokens for each token they contributed. Participants were told that a monetary incentives policy would be implemented via a computer program 2–5 days after the experiment. In the context of the social incentives, participants were told that their behaviors would be evaluated and judged by another group of participants who did not join in the PGG. Based on their performance, participants would receive positive (e.g., you are a generous person) or negative feedback (e.g., you are a selfish person). The participants were told that their behaviors would be evaluated by another group of individuals, rather than by the participants in the PGG, in order to simplify the relationships within the group. The participants were told that another group of people would only be aware of their performance on the task, without any personal information about the participants being revealed. Participants received an email with comments about their choices in the social incentives context 2–5 days after the experiment. The comments were given by five participants who did not participate in the PGG. Finally, an impunity context was used as the control condition, in which participants were instructed that no additional rules would be implemented. The three conditions were presented in a pseudorandom manner to participants, with 20 rounds for each condition, resulting in 60 rounds in total.

In each round, every player was given an endowment of 10 tokens. They then had to decide how many tokens to keep for themselves and how many to contribute to the public pool. Each token that a person kept for themselves resulted in one money unit to that person. Each token contributed to the public pool was multiplied by 1.6 and then divided equally among the four group members. All participants played the PGG individually and were not told about the outcomes or choices of other players. Prior to the PGG, individuals were asked to rate their first and second-order beliefs on an 11-point Likert scale, with 0 tokens representing the lowest possible score and 10 tokens being the highest. The first-order belief is the participants' expectation of other players' contributions. They answer the question, "How many tokens do you think other players will contribute to the group project, on average?" for each context. The second-order belief is the participants' belief about other players' expectations of participants' contributions; they answer the question "How many tokens do you think other players believe that you will contribute to the group project?" The self-ratings scores were used to measure participants' expectations, which is consistent with previous studies (Chang et al., 2011; Gong et al., 2019).

Experimental procedure

The experiment procedure is illustrated in Figure 1A. Participants sat in a quiet room to complete the screening and read the task instructions. Before the PGG, participants completed an LSRP scale. After PGG instructions, participants reported their first- and second-order beliefs about other players during the game. To avoid social desirability bias during the game, the PGG was always referred to as the "interaction game." Participants were instructed that they would play a game with anonymous players across multiple rounds, and that the players were played by different people in each round. Participants were paid based on their performance by randomly selecting several rounds to calculate their earnings at the end of the game. All participants underwent several practice rounds to ensure they fully understood the task before the task. Stimulus presentation and behavioral data collection were implemented using Psychtoolbox (http://psychtoolbox.org/; Brainard & Vision, 1997) in MATLAB (Version R2017a).

Statistical analysis

All statistical analyses were conducted in R Studio 4.2.1 (Ihaka & Gentleman, 1996).

Linear mixed models (LMMs)

Linear mixed models were estimated in "lmerTest" package (Kuznetsova et al., 2017), which provides p values for tests for fixed effects and implements the Satterthwaite's method for approximating degrees of freedom for the t and F tests. For significant fixed effects, the "emmeans" package was used to test post hoc contrasts. For the interaction between categorical and continuous variables, we compared the slopes of the fitted lines. Bonferroni correction was used to control for type I error inflation caused by multiple comparisons.

LMM1: contribution is the dependent variable; fixed effects include an intercept, the main effect of primary psychopathic traits (continuing variable), the main effect of secondary psychopathic traits (continuing variable), the main effect of context (categorical variable), the interaction of primary psychopathic traits and context, and the interaction of second psychopathic traits and context; the random effect is the random participant intercept.

LMM2: the LMM2 is similar to LMM1 except that the first-order belief rating score is the dependent variable.

LMM3: the LMM3 is similar to LMM1 except that the second-order belief rating score is the dependent variable.

LMM4: the LMM4 is similar to LMM1 except that guilt aversion is the dependent variable.

Computational modeling

Model description

Model 1: Self-collective interests integration model

This model assumes that participants make decisions by integrating the interests of the self and the interests from a collective perspective (i.e., the interests of all players). This model was initially adapted from Park et al. (2011) (see Supplementary section 1 in Appendix S1).

Model 2: Guilt aversion model

The guilt aversion model assumes that people are motivated not only by self-interest, but also by a desire to avoid the guilt that comes from disappointing others (Chang et al., 2011; Gong et al., 2019). The utility function of the guilt aversion model posits that individuals will avoid decisions that may result in feelings of guilt:

$$SV_{C_s} = \text{self} - \omega \cdot \text{anticipated guilt} \quad (1)$$

and

$$\text{anticipated guilt} = E_2C_s - C_s, \quad (2)$$

where $SV_{C_s}\omega_{\text{social}}$ is the participants' subjective value and *self* denotes participants' self-interests. The term "subjective value" is used because participants' expected value includes a non-material component in addition to the material reward, namely guilt aversion. The $C_s \in (0,10)$ denotes the self-contribution amount, and E_2C_s denotes participants' expectations about other players' expectations of participants' average contribution amount, that is, the second-reported first-order rating score. The ω is a free parameter that measures the magnitude of guilt aversion ($0 \le \omega \le 1$). We test the range of ω value by parameter recovery and this range can be recovered well (see Supplementary section 1 in Appendix S1). Large ω indicates more guilt aversion and ω approach to zero denotes no guilt aversion. Guilt aversion was estimated in a different context that allowed us to test whether participants change the levels of guilt aversion across contexts. Parameters ω_{monetary}, ω_{social} and ω_{control} denote the guilt aversion in the monetary incentives context, social incentives context, and control condition, respectively. The self-interest is the participant's expected payoff of himself/herself and calculated as follows:

$$\text{self} = 10 - C_s + \frac{1.6(C_s + 3 \cdot E_1C_o)}{4}, \quad (3)$$

where E_1C_o denotes participants' expectations about the average contribution amount of other players, that is, the self-reported first-order rating score. A participant's expected payoff is equal to their original endowment minus their contribution to the public pool, plus an expected reward, $\frac{1.6(C_s+3 \cdot E_1C_o)}{4}$. The expected reward is calculated by taking into account the participant's contribution as well as the contributions of other players. The other player's contributions were denoted by participants' expectations about the average contribution of other players, that is, E_1C_o.

The probability of choosing each choice was modeled using a softmax function:

$$p_{C_s} = \frac{e^{\beta \cdot SV_{C_s}}}{\sum_{C_s=0}^{10} e^{\beta \cdot SV_{C_s}}}, \quad (4)$$

where β denotes to the inverse softmax temperature ($0 \le \beta \le 1$), which measures the sensitivity of an individual's choice to the difference in utilities between options of choice and unchosen. In this function, the numerator represents the subjective value of the selected option, and the denominator is the sum of the subjective values of all options. The unknown parameters of the above models were estimated dependently for each context. Parameters for the above models were estimated in MATLAB (Versions of R2016b) using the fmincon optimization function for each participant by maximizing the log-likelihood (LL) for each participant as the following function:

$$LL = \max\left(\sum_{t=1}^{60}\left(p_{C_{s(t)}}|\theta|\right)\right), \quad (5)$$

where $p_{Cs(t)}$ denotes the probability of participants choosing C_s to the public pool in t round, and θ denotes the free parameters, including β, ω_{monetary}, ω_{social} and ω_{control} in the guilt aversion model. We reduced the likelihood of the model converging on a local minimum by selecting 200 random start locations.

Model 3: The combined model

This model combined the computation components of Model 1 and Model 2 (see Supplementary section 1 in Appendix S1).

Model comparisons

We compared three models by calculating the corrected Akaike information criterion for small sample (AICc; Hurvich & Tsai, 1989). The lower value of Bayesian information criterion with a correction for sample sizes (BICc) and Akaike information criterion with a correction for sample sizes (AICc) averaged across participants indicates a better fit. Protected exceedance probability (PEP) was used to examine which model can describe our data better than the alternative model at the group level (Rigoux et al., 2014). The analysis results from the winning model are reported in the present study. The mean value and 95% confidence interval (CI) of the parameters are reported in Table S1. See Supplementary section 1 in Appendix S1 for the analysis of parameter recovery, and model validation of the winning model, which ensures that each component of the model can be well dissociated and identified and that our model predicts the real behaviors very well.

RESULTS

Model-free results

Contribution

A linear mixed regression model (LMM1, see Methods) was built with the independent variables of primary psychopathic traits, secondary psychopathic traits, and context (a categorical variable: social incentives, monetary incentives, and a control

condition) to predict participants' contributions. We found that the interaction of context and primary psychopathic traits (F[2, 114] = 5.595, p < .01) was significant. The primary psychopathic traits significantly predicted participants' contribution in the social incentive context (t[114] = −2.601, $p_{corrected}$ < .05), but not in the monetary incentive context (t[114] = 0.722, $p_{corrected}$ = .472), or the control context (t[114] = −1.294, $p_{corrected}$ = .202). The post hoc comparison showed that the slope of the primary psychopathic traits on predicting contribution in the social incentives context (beta = −2.472, 95% CI [−4.731, –0.214]) was significantly smaller than that in monetary context (beta = −0.583, 95% CI [−1.680, 2.841]; t[114] = 3.293, $p_{corrected}$ < .01). The slope of primary psychopathic traits on predicting contribution in the control condition (beta = −1.431, 95% CI [−3.693, 0.830]) did not significantly differ from that in the social incentive context (t[114] = 1.121, $p_{corrected}$ = .500) or the monetary incentive context (t[114] = −2.172, $p_{corrected}$ = .081; see Figure 2A). The main effects of primary psychopathic traits (F[1, 57] = 2.108, p = .153), secondary psychopathic traits (F[1, 57] = 0.031, p = .848), and context (F[2, 114] = 2.212, p = .109) were not significant (Figure S1A). The interaction of secondary psychopathic traits and context was not significant (F[2, 114] = 0.412, p = .656). These results suggest that individuals with higher levels of primary psychopathic traits were less likely to cooperate only in the social incentives context, but not in the monetary incentive context and/or control context.

Self-reported beliefs

A linear mixed regression model (LMM2; see Methods) was built with the independent variables of primary psychopathic traits, secondary psychopathic traits, and the context to predict participants' expectation about others' contribution (i.e., the first-order belief). We found that the main effect of context significantly predicted participants' expectation about others' contribution (F[1, 57] = 5.592, p < .01; post hoc comparison: monetary incentives [M = 7.228, SD = 0.307] - social incentives [M = 4.879, SD = 0.306], t[114] = 2.341, $p_{corrected}$ < .001; monetary incentive – control [M = 3.163, SD = 0.308], t[114] = 4.070, $p_{corrected}$ < .001; social incentives – control, t[114] = 1.718, $p_{corrected}$ < .001). The main effects of primary psychopathic traits (F[1, 57] = 0.943, p = .327) and secondary psychopathic traits (F[1, 57] = 0.172, p = .674) were not significant. The interaction of the primary psychopathic traits with context (F[2, 114] = 1.701, p = .193), and the interaction of secondary psychopathic traits with context were not significant (F[2, 114] = 0.327, p = .708; see Figure 2B).

A linear mixed regression model (LMM3, see Methods) was built with the independent variables of primary psychopathic traits, secondary psychopathic traits, and context to predict a participant's inferred other player's expectations about their own decisions (i.e., second-order belief). We found that none of the main effects or interaction can predict a participant's inferred other player's expectations about their own decisions: the main effect of primary psychopathic traits (F[1, 57] = 0.778, p = .385); the main effect of secondary psychopathic traits (F[1, 57] = 0.011, p = .933); the main effect of context (F[2, 114] = 2.179, p = .118); the interaction of primary psychopathic traits with context (F[2, 114] = 0.242, p = .775); and the interaction of secondary psychopathic traits with context (F[2, 114] = 0.585, p = .552; see Figure 2c and Figure S1B,C). These findings suggested that the expectations about other players' contribution of individuals with different levels of primary psychopathic traits and secondary psychopathic traits were not significantly affected.

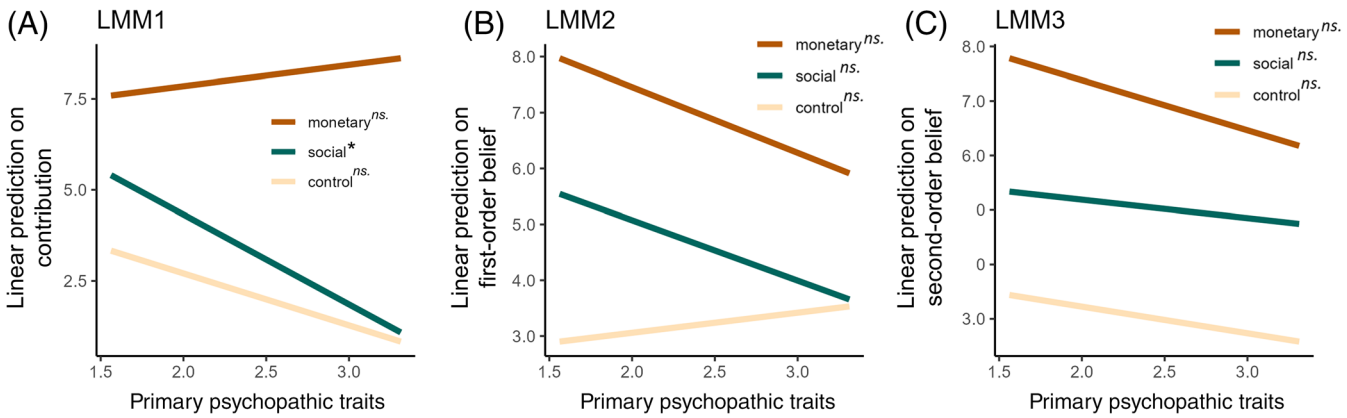

FIGURE 2 Model-free analysis results. The interaction of primary psychopathic traits and context on (A) contribution, (B) self-reported score of first-order belief, and (C) self-reported score of second-order belief. The first-order belief is participants' expectation about other players' contribution. Specifically, participants were asked "How many tokens do you think other players will contribute to the group project, on average?" The second-order belief is participants' inferred expectations of their own contribution from other players' perspectives. Participants were asked "How do you think other players will rate your contribution to the group project?" Neither the first-order belief nor the second-order belief was found to be predicted by the score on primary psychopathic traits in all contexts. One dot represents one participant. *ns.*, * $p_{corrected}$ < .05.

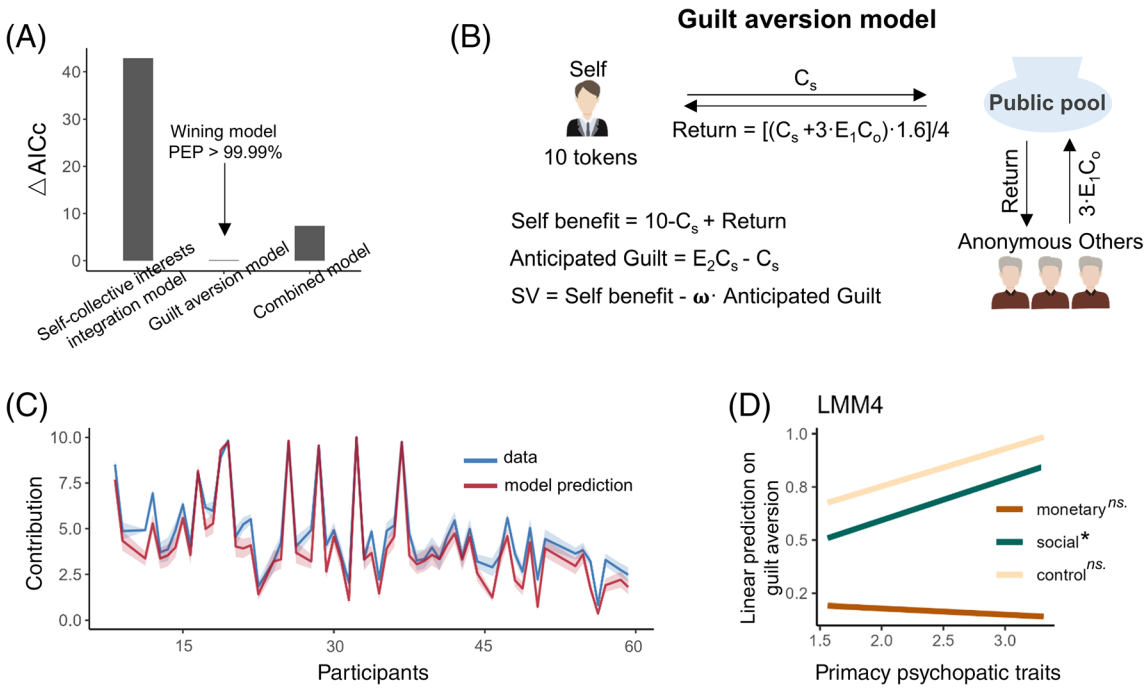

FIGURE 3 Model-based analysis results. (A) Model comparison AICc showed that participants' decisions were best described by Model 2 (guilt aversion model). (B) The mathematical framework of the guilt aversion model. (C) Model prediction. The average contribution of participants is denoted by the blue line, while the predictions of the guilt aversion model are denoted by the red line. The shaded area in the graph represents the standard error (*SE*) of the data (see Figure S2 for the predictions in each context and the predictions of the alternative model). (D) The interaction of primary psychopathic traits and context on guilt aversion. *Ns.*, $^{*}p_{corrected} < .05$.

Model-based results

The results showed that Model 2, the guilt aversion model, had the lowest BICc and AICc scores (see Figure 3A and Table S1). The PEP indicates that Model 2 explains our data better than alternative models (see Table S1). The guilt aversion model assumed that people are motivated not only by self-interest, but also by a desire to avoid the guilt that comes from disappointing others (see Figure 3B). We compared the goodness-of-fit of models by the AICc and PEP and found the guilt aversion model (Model 2) fitted participants' decisions best (Figure 3A). The prediction of the guilt aversion model agreed well with the observed data (Figure 3C). The prediction of the guilt aversion model for each context and the prediction of another alternative model are reported in Figure S2. We next conducted a linear regression mixed model (LMM4; see Methods) with the independent variables of primary psychopathic traits, secondary psychopathic traits, and context to predict participants' guilt aversion. We found that the interaction of primary psychopathic traits and context significantly predicted participants' guilt aversion ($F[2, 114] = 3.059$, $p = .050$). The primary psychopathic traits significantly predicted participants' guilt aversion in the social incentive context ($t[114] = -2.072$, $p_{corrected} < .05$), but not in the monetary incentive context ($t[114] = 0.717$, $p = .471$), or the control context ($t[114] = 1.764$, $p_{corrected} = .082$). The post hoc comparison showed that the slope of primary psychopathic traits on predicting guilt aversion in the social incentives context (beta = 0.19, 95% CI [−0.050, 0.442]) was marginally significantly larger than that in the monetary incentives context (beta = −0.032, 95% CI [−0.271, 2.214]; $t[114] = 2.208$, $p_{corrected} = .072$). The slope of primary psychopathic traits on predicting contribution in the control condition (beta = 0.17, 95% CI [−0.063, 0.418]) did not significantly differ from that in the social incentive context ($t[114] = -0.162$, $p_{corrected} = .977$) or the monetary incentive context ($t[114] = 2.052$, $p_{corrected} = .102$, see Figure 3D). The main effects of primary psychopathic traits ($F[1, 57] = 1.917$, $p = .171$), secondary psychopathic traits ($F[1, 57] = 0.014$, $p = .992$), and context ($F[2, 114] = 0.947$, $p = .382$) were not significant. The interaction of secondary psychopathic traits and context was not significant ($F[2, 114] = 0.886$, $p = .412$, see Figure S1D). These findings suggest that in the social incentive context, individuals with higher levels of primary psychopathic traits were more averse to guilt caused by disappointing other players. This effect was not observed in the monetary incentive context or the control context.

DISCUSSION

The present study examined the effects of social incentives and monetary incentives on cooperative behaviors and their associations with non-clinical psychopathic personality traits. Healthy adults with varying psychopathic traits played a one-shot public goods game with anonymous players in a monetary incentives

context where decisions may have resulted in making or losing money, a social incentives context where decisions would be judged by others, and a control context where no additional incentives were implemented. We found that both social and monetary incentives were effective in promoting cooperative behaviors, compared to the control condition. Participants in a monetary incentive context overall improved cooperative behaviors; this effect was not correlated with participants' psychopathic traits. Social incentives also improved cooperative behaviors, but not as strongly as monetary incentives. We found that social incentives have a differential effect on cooperation depending on an individual's level of primary psychopathic traits. Those with higher primary psychopathic traits were found to be less likely to cooperate in the context of the social incentives. The results indicated that the cooperative behaviors of participants with higher primary psychopathic traits were less influenced by social incentives. We found that none of the contexts had an effect on cooperative behavior that could be predicted by secondary psychopathic traits. The findings from the study showed that social incentives could foster cooperative behaviors, an effect that diminishes with primary psychopathic traits.

We further investigated whether participants with high psychopathic traits could appropriately build social expectations or/and adjust their behavior responses to match others' expectations. Our Hypothesis 1 is that if participants with high psychopathic traits have deficits in building appropriate social expectations, the participants' self-reported scores about social expectations should be predicted by the scores on psychopathic traits. The results showed that psychopathic traits cannot predict either participants' expectations of others (first-order beliefs) or the inferred others' expectations (second-order beliefs). That is, individuals with higher psychopathic traits did not have different social expectations from individuals with lower psychopathic traits. Our findings suggested that individuals with high psychopathic traits might not be deficient in forming appropriate social expectations.

Our Hypothesis 2 was that if participants with high psychopathic traits were reluctant to adjust their behaviors to match others' expectations, a negative association between psychopathic traits and guilt aversion would be observed. Guilt aversion reflects the degree to which participants wished to avoid the anticipated guilt that comes from disappointing other players. Higher guilt aversion indicates participants were more likely to adjust their behaviors to match others' expectations and to avoid guilt. We found that higher primary psychopathic traits were correlated with lower guilt aversion only in the social incentive context. These findings indicated that participants with higher primary psychopathic traits are reluctant to adjust their behaviors to match others' expectations in the context where their decisions would be observed and judged by others. However, in the context where their decisions could result in a monetary gain or loss, those with higher levels of psychopathic traits did not differ from those with lower psychopathic traits in terms of guilt aversion. Our results are not consistent with previous studies that suggested that psychopathic personality may involve deficits in mentalizing ability and difficulty following social expectations (Blair, 2007; Harenski et al., 2010; Koenigs et al., 2012), probably because of the difference in research sample. The current study and the study of Gong et al. (2019) recruited healthy adults as a research sample. The findings of the current study are consistent with those of Gong et al. (2019), which found that individuals with higher scores on psychopathic traits repaid less money to their investors in the trust game, indicating an effect caused by participants' diminished guilt aversion associated with disappointing the investor rather than participants' distortions in building social expectations (Gong et al., 2019). Together, we illustrated that individuals with higher psychopathic traits tended to cooperate less in social incentive contexts not due to their deficits in building appropriate social expectations, but because they experience less guilt when they violate social expectations.

We did not observe any behavioral variable that was correlated with secondary psychopathic traits, which is consistent with previous studies (Gervais et al., 2013; Rilling et al., 2007; White, 2014) and indicated a specific role of primary psychopathic traits in cooperative decision-making. The effect of monetary incentives on promoting cooperative behaviors was consistent with prior studies (Noussair & Tucker, 2005; Weber et al., 2018), suggesting that money can be a powerful incentive to increase cooperation in general. The main characteristic of psychopathic personalities is the interpersonal–affective disturbance, including the lack of remorse, empathy, and guilt (Blair, 1995; Hare, 2003; Hare et al., 1991). In line with this view, we found that individuals with higher psychopathic traits had less guilt aversion when they did not follow the social expectations to make a decision; this effect was only observed in the context of the social incentives. In the monetary incentive context and control context, individuals with different levels of psychopathic traits did not show different cooperative behaviors. The unique effect of social incentives suggested a flexible psychological mechanism in individuals with psychopathic personalities to adapt to social contexts, with a goal to get as much financial benefit as possible. A previous study found that students majoring in commerce and business scored significantly higher on primary psychopathic personality than students studying in noncommercial majors (Wilson & McCarthy, 2011). These findings highlight that subclinical primary psychopaths have a money-seeking behavioral style and may flexibly exploit others in specific situations (Gervais et al., 2013). One limitation of the present study is that all of the participants were university students, which may limit the generalizability of the findings.

CONCLUSION

The effect of social incentives on promoting cooperative behaviors depends on primary psychopathic traits: individuals with higher primary psychopathic traits were less likely to increase cooperation in the social incentives context where their decisions would be judged by others. This was caused by participants' diminished guilt aversion when they intentionally violated social expectations, rather than by their distortion in building appropriate social expectations.

AUTHOR CONTRIBUTIONS

Xiaoyan Wu and Xu Gong: Conceptualization. Xiaoyan Wu and Ruida Zhu: Methodology, Writing – review & editing. Xiaoyan Wu: Data collection, Data Analysis, Visualization, Writing-original draft preparation. Chao Liu: Supervision, Project administration, Funding acquisition. Yuejia Luo: Project administration, Funding acquisition.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (32271092 and 32130045), the Major Project of National Social Science Foundation (19ZDA363). Thanks to Lu Yan for her assistance in collecting data and Zhuo Zhang of the China University of Political Science and Law for providing the Chinese version of the LSRP.

ETHICS STATEMENT

The procedures used in this study adhere to the tenets of the Declaration of Helsinki. This study and recruitment of subjects were approved by the Ethics Committee of Beijing Normal University (No. CNL-A-0001-009).

ORCID

Xiaoyan Wu https://orcid.org/0000-0002-2683-3807

REFERENCES

Anderson, D. A. (1999). The aggregate burden of crime. *The Journal of Law Economics*, *42*(2), 611–642. https://doi.org/10.2139/ssrn.147911

Blair, R. J. R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, *11*(9), 387–392. https://doi.org/10.1016/j.tics.2007.07.003

Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, *57*(1), 1–29. https://doi.org/10.1016/0010-0277(95)00676-P

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, *70*(3), 560–572. https://doi.org/10.1016/j.neuron.2011.02.056

Cima, M., & Raine, A. (2009). Distinct characteristics of psychopathy relate to different subtypes of aggression. *Personality and Individual Differences*, *47*(8), 835–840. https://doi.org/10.1016/j.paid.2009.06.031

Curry, O., Chesters, M. J., Viding, E. J. P., & Differences, I. (2011). The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. *Personality and Individual Differences*, *50*(6), 804–809. https://doi.org/10.1016/j.paid.2010.12.036

Falkenbach, D., Poythress, N., Falki, M., & Manchak, S. (2007). Reliability and validity of two self-report measures of psychopathy. *Assessment*, *14*(4), 341–350. https://doi.org/10.1177/1073191107305612

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

Gervais, M. M., Kline, M., Ludmer, M., George, R., & Manson, J. H. (2013). The strategy of psychopathy: Primary psychopathic traits predict defection on low-value relationships. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1757), 20122773. https://doi.org/10.1098/rspb.2012.2773

Glenn, A. L., Kurzban, R., & Raine, A. (2011). Evolutionary theory and psychopathy. *Aggression and Violent Behavior*, *16*(5), 371–380. https://doi.org/10.1016/j.avb.2011.03.009

Gong, X., Brazil, I. A., Chang, L. J., & Sanfey, A. G. (2019). Psychopathic traits are related to diminished guilt aversion and reduced trustworthiness during social decision-making. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-43727-0

Hare, R. D. (1998). Psychopaths and their nature: Implications for the mental health and criminal justice systems. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, Criminal, and Violent Behavior* (pp. 188–212). The Guilford Press. https://doi.org/10.4324/9781003193456-1

Hare, R. D. (2003). *The psychopathy checklist–revised* (p. 412). Multi-HealthSystems.

Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, *100*(3), 391–398. https://doi.org/10.1037/0021-843X.100.3.391

Harenski, C. L., Harenski, K. A., Shane, M. S., & Kiehl, K. A. (2010). Aberrant neural processing of moral violations in criminal psychopaths. *Journal of Abnormal Psychology*, *119*(4), 863–874. https://doi.org/10.1037/a0020979

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297–307. https://doi.org/10.1093/biomet/76.2.297

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational Graphical Statistics*, *5*(3), 299–314. https://doi.org/10.2307/1390807

Johnston, L., Hawes, D., & Straiton, M. (2014). Psychopathic traits and social cooperation in the context of emotional feedback. *Psychiatry, Psychology and Law*, *21*(5), 767–778. https://doi.org/10.1080/13218719.2014.893550

Koenigs, M., Kruepke, M., & Newman, J. (2010). Economic decision-making in psychopathy: A comparison with ventromedial prefrontal lesion patients. *Neuropsychologia*, *48*(7), 2198–2204. https://doi.org/10.1016/j.neuropsychologia.2010.04.012

Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive Affective Neuroscience*, *7*(6), 708–714. https://doi.org/10.1093/scan/nsr048

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lee, K., & Ashton, M. C. (2005). Psychopathy, Machiavellianism, and narcissism in the five-factor model and the HEXACO model of personality structure. *Personality and Individual Differences*, *38*(7), 1571–1582. https://doi.org/10.1016/j.paid.2004.09.016

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral Brain Sciences*, *18*(3), 523–541. https://doi.org/10.1017/S0140525X00039595

Mokros, A., Menner, B., Eisenbarth, H., Alpers, G. W., Lange, K. W., & Osterheider, M. (2008). Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *Journal of Abnormal Psychology*, *117*(2), 406–413. https://doi.org/10.1037/0021-843X.117.2.406

Noussair, C., & Tucker, S. (2005). Combining monetary and social sanctions to promote cooperation. *Economic Inquiry*, *43*(3), 649–660. https://doi.org/10.1093/ei/cbi045

Park, S. Q., Kahnt, T., Rieskamp, J., & Heekeren, H. R. (2011). Neurobiology of value integration: When value impacts valuation. *Journal of Neuroscience*, *31*(25), 9307–9314. https://doi.org/10.1523/JNEUROSCI.4973-10.2011

Park, S. A., Sestito, M., Boorman, E. D., & Dreher, J. C. (2019). Neural computations underlying strategic social decision-making in groups. *Nature Communications*, *10*(1), 1–12. https://doi.org/10.1038/s41467-019-12937-5

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies—Revisited. *Neuroimage*, *84*, 971–985. https://doi.org/10.1016/j.neuroimage.2013.08.065

Rilling, J., Glenn, A., Jairam, M., Pagnoni, G., Goldsmith, D., Elfenbein, H., & Lilienfeld, S. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological*

Psychiatry, *61*(11), 1260–1271. https://doi.org/10.1016/j.biopsych.2006.07.021

Skeem, J., Johansson, P., Andershed, H., Kerr, M., & Louden, J. E. (2007). Two subtypes of psychopathic violent offenders that parallel primary and secondary variants. *Journal of Abnormal Psychology*, *116*(2), 395–409. https://doi.org/10.1037/0021-843X.116.2.395

ten Brinke, L., Black, P. J., Porter, S., & Carney, D. R. (2015). Psychopathic personality traits predict competitive wins and cooperative losses in negotiation. *Personality Individual Differences*, *79*, 116–122. https://doi.org/10.1016/j.paid.2015.02.001

Verma, J., & Verma, P. (2020). Determining sample size in general linear models. In *Determining sample size and power in research studies* (pp. 89–119). Springer. https://doi.org/10.1007/978-981-15-5204-5_7

Weber, T. O., Weisel, O., & Gächter, S. (2018). Dispositional free riders do not free ride on punishment. *Nature Communications*, *9*(1), 1–9. https://doi.org/10.1038/s41467-018-04775-8

White, B. A. (2014). Who cares when nobody is watching? Psychopathic traits and empathy in prosocial behaviors. *Personality Individual Differences*, *56*, 116–121. https://doi.org/10.1016/j.paid.2013.08.033

Widom, C. S. (1976). Interpersonal conflict and cooperation in psychopaths. *Journal of Abnormal Psychology*, *85*(3), 330–334. https://doi.org/10.1037/0021-843X.85.3.330

Wilson, M. S., & McCarthy, K. (2011). Greed is good? Student disciplinary choice and self-reported psychopathy. *Personality Individual Differences*, *51*(7), 873–876. https://doi.org/10.1016/j.paid.2011.07.028

Wu, J., Balliet, D., & Van Lange, P. A. (2016). Gossip versus punishment: The efficiency of reputation to promote and maintain cooperation. *Scientific Reports*, *6*(23), 919. https://doi.org/10.1038/srep23919

Zheng, C. (2011). Research on the emotional mechanism of adolescents with psychopathic traits. Master's thesis, China University of Political Science and Law.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wu, X., Zhu, R., Gong, X., Luo, Y., & Liu, C. (2023). Social incentives foster cooperation through guilt aversion: An effect that diminishes with primary psychopathic traits. *PsyCh Journal*, *12*(3), 389–398. https://doi.org/10.1002/pchj.641