

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/128115>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

© 2019 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

Cruel to be Kind: Factors Underlying Altruistic Interpersonal Affect Worsening

Belén López-Pérez, Laura Howells, & Michaela Gummerum

University of Plymouth

Correspondence concerning this article should be addressed to Belén López-Pérez, School of Psychology, Cognition Institute, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK. Electronic mail may be sent to belen.lopez-perez@plymouth.ac.uk

This is the peer-reviewed version of the following article: López-Pérez, B., Howells, L., & Gummerum, M. (in press). Cruel to be Kind: Factors Underlying Altruistic Interpersonal Affect Worsening. *Psychological Science*. This article may not exactly replicate the final version published. It is not the copy of record.

Abstract

When aiming to improve another person's long-term well-being, people may choose to induce that person to experience a negative emotion in the short term. We labelled this form of agent–target interpersonal emotion regulation *altruistic affect worsening* and hypothesized that it may happen when three conditions are met: (1) The agents experience empathic concern for the target of the affect-worsening process; (2) the negative emotion to be induced helps the target achieve a goal (anger for confrontation or fear for avoidance); and (3) there is no benefit for the agent. This hypothesis was tested by manipulating perspective taking instructions and the goal to be achieved whilst participants ($N = 140$) played a computer video game with different goals. Participants following other-oriented perspective taking instructions decided to induce more anger or fear in a supposed fellow participant working to achieve a confrontational or avoidance goal, respectively.

Keywords: altruistic affect worsening; interpersonal emotion regulation; goal; perspective taking; emotion.

Cruel to Be Kind: Factors Underlying Altruistic Interpersonal Affect Worsening

Sometimes people cause a loved one to experience a negative emotional state if they think that this will increase that other person's (long-term) well-being. But why would feeling bad be beneficial? According to an instrumental approach to emotion regulation, people may choose to feel a positive or negative emotion in the short term if doing so maximizes the attainment of a specific long-term goal (Erber & Erber, 2000; Tamir, 2009; Tamir & Ford, 2009). For example, people may choose to feel anger when pursuing confrontation goals (e.g., dealing with someone who cheated) or fear when pursuing avoidance goals (e.g., escaping from a scary situation) because these negative emotions are seen as beneficial for achieving these specific goals (Tamir & Ford, 2009; Tamir, Mitchell, & Gross, 2008).

Research on the regulation of others' emotions (i.e., interpersonal emotion regulation; Gross & Thompson, 2007) has for a long time followed a hedonic approach that suggests people may attempt to decrease others' negative emotions if those emotions are perceived as harmful for the others (Zaki & Williams, 2013). However, people may also engage in instrumental affect worsening when regulating others' emotions: An agent may choose to make a target feel bad (1) if this negative emotion allows the target to achieve a goal and (2) if the agent him- or herself can benefit from this interpersonal emotion regulation by obtaining a desirable outcome (Netzer, Van Kleef, & Tamir, 2015). In this case, instrumental interpersonal affect worsening would be purely egoistically motivated. But would people choose to change others' emotions for altruistic reasons? Or, put differently, would an agent make others feel bad in the short term if this negative affect entails a benefit solely for the target of the regulation process and not for the agent him- or herself? The aim of this study was to investigate conditions for such *altruistic affect worsening*.

We suggest that three conditions must be met for altruistic affect worsening to happen. First, the agent's motivation has to be altruistic, that is, the final aim of the agent's

action is to *increase the target's well-being* rather than to obtain a personal benefit or goal, according to the classic definition of altruistic motivation (Batson, 2011). Second, the agent must aim to instil a negative emotion in the target that is *beneficial for the target's goal pursuit* (e.g., making the target feel anger to achieve confrontation goals or fear to achieve avoidance goals; Netzer et al., 2015; Tamir & Ford, 2009). In these situations affect worsening is seen as a means to an end, not an end itself (Niven, Totterdell, & Holman, 2009; Tamir, 2015). Third, altruistic motivation and altruistic affect worsening are more likely when the agent *empathizes with the target*.

To test these claims we reconciled two different research traditions. First, we drew on Batson and colleagues' (e.g., Batson, 2011) experimental methods to manipulate empathic concern through perspective taking instructions. People who received other-oriented perspective taking instructions (e.g., "imagine how the other person is feeling in a certain situation") were shown to be more likely to experience empathic concern and to act altruistically than people who received objective perspective taking instructions (Batson, Early, & Salvarani, 1997). So far, this line of research has focussed on behaviour alone as a means to increase others' benefits or decrease others' suffering (e.g., taking electric shocks on behalf of another; Batson, Duncan, Ackerman, Buckley, & Birch, 1981). However, it has not investigated whether empathic concern leads people to use emotions to benefit others (e.g., to help them achieve a specific goal). Secondly, we relied on Tamir and colleagues' (e.g., Tamir, 2015) procedures to study people's explicit and implicit emotional preferences for others and perceptions of emotion utility. Although this research has shown that people may choose to make others feel bad if they themselves benefit (Netzer et al., 2015), it has not assessed whether people engage in affect worsening for the sole benefit of another. Thus, we sought not only to extend previous research findings by relying on reliable experimental designs but also to bridge different traditions to expand current knowledge on interpersonal

emotion regulation. This may enhance our understanding of the dynamics of social interaction and social cognition by providing further information about adults' emotion–outcome expectancies and how they balance (emotional) costs and benefits when regulating other people's emotions.

To study altruistic affect worsening, we focussed on the process model of emotion regulation (Gross, 2007), which posits that people may change their own and others' emotions by selecting a strategy that influences a particular stage of the emotion process. People may change emotions by selecting or modifying a situation (e.g., not going to a party), diverting attention (e.g., looking away), changing what they think about the situation (e.g., reappraisal), or altering their physiological response (e.g., suppression). We focussed on the strategy of *situation selection*, which involves selecting or avoiding a stimulus or a situation in order to experience a specific emotion (Gross, 2007). Previous research (e.g., Netzer et al., 2015; Tamir et al., 2008) has shown how participants selected different emotion-inducing stimuli to change the emotional experience in themselves or others to attain specific goals.

We hypothesized that participants would show altruistic affect worsening under the conditions outlined above. Specifically, compared to participants in an objective perspective taking condition, participants in an other-oriented perspective taking condition should select for a target (a) more negative emotional stimuli at the risk of lowering their own chances of earning £50 (*empathy hypothesis*) and (b) negative stimuli consistent with the target's goal. That is, angry emotional stimuli should be chosen in a confrontation goal condition and fearful emotional stimuli in an avoidance goal condition (*beneficial goal hypothesis*). Furthermore, participants in an other-oriented perspective taking condition (c) should rate their chosen emotional stimuli as more beneficial than the other stimuli in pursuing the target's goals (at their own expense; *altruistic motivation hypothesis*).

Method

Participants

One hundred and forty adults (86 women, 54 men; $M_{\text{age}} = 30.85$ years, $SD = 13.68$; age range: 18–71 years) were recruited from a paid pool at the authors' institution and completed the study in exchange for payment (£4/\$6). An a priori power analysis showed that 35 participants per condition should have 80% power to detect an effect size (f) of 0.50.

Design

A 2×2 between-subjects design was employed with the two independent factors perspective taking (other-oriented, objective) and goal (confrontation, avoidance). Participants were randomly allocated to one of four conditions: (1) other-oriented perspective taking/confrontation; (2) objective perspective taking/confrontation; (3) other-oriented perspective taking/avoidance; (4) objective perspective taking/avoidance.

Procedure

The study was presented as an examination of performance in one of two computer-based video games. Participants were tested in groups of four. Closeness between participants was controlled by making sure participants did not know each other. Each participant completed the study in a separate cubicle. After signing the consent form, participants rated their current mood on the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988). Then, as a cover story, participants were told that they would be paired with another anonymous person. If assigned to the role of Player A, they had to write down a personal statement so that Player B could get to know them before making choices for them in the game. If assigned to the role of Player B, they would receive a personal statement from Player A before making choices for him/her in the game. In fact, all participants were allocated to the role of Player B.

Participants were told that prior to reading Player A's note they would receive instructions to make sure they all had a similar emotional experience. Participants in the

other-oriented perspective taking condition were asked to imagine how the other player would feel in the described scenario, whereas participants in the objective perspective taking condition were told to remain detached about the note (see Batson et al., 2007). As we aimed to test whether affect worsening may happen for altruistic reasons and given that an altruistic motivation is more likely to happen when empathic concern is experienced (Batson, 2011), we decided to include only those manipulations that would most likely isolate the experience of empathic concern from the experience of personal distress (i.e., other-oriented vs. objective perspective taking instructions; see Batson et al., 1988). Therefore, we decided not to include a no-instruction condition because (a) previous research has shown that when not given any instructions, people tend to take another's perspective (see Batson, 2011, for a review) and (b) a no-instruction condition can increase both empathic concern and personal distress, as these emotions usually co-occur (e.g., Barraza & Zak, 2009). Hence, in a no-instruction condition participants might experience empathic concern and personal distress which might entail both an altruistic and egoistic motivation (Batson, 2011).

Next, participants received a sealed envelope with a purported handwritten communication from Player A that described Player A's recent break-up and how upset and helpless Player A was feeling about it (taken from Batson et al., 2007), in order to provoke empathic concern in the participants. After reading Player A's note, participants rated how they felt towards Player A using the Empathic Response Questionnaire (Batson, Fultz, & Schoenrade, 1987).

In addition, participants were tested in one of two goal-pursuit conditions. They were asked to play different games so we could manipulate the goal to be achieved. In the confrontation goal condition, the actual participants (as well as the supposed partner) were asked to play the game 'Soldier of Fortune', a first-person shooter game with a clear confrontation goal (i.e., to kill as many enemies as possible; see Netzer et al., 2015; Tamir et

al., 2008). In the avoidance goal condition, participants were asked to play the game ‘Escape Dead Island’, a first-person game with the goal of avoidance (i.e., escaping from a room without being killed by zombies). We chose this game to test fear preference for achieving an avoidance goal, as previous literature has extensively linked avoidance behaviour with fear (Carver, 2001; Frijda, 1986; Ohman, 1993). For both games, participants were informed that depending on their own individual performance (i.e., number of individuals killed in ‘Soldier of Fortune’ and distance travelled in ‘Escape Dead Island’) they would receive a number of raffle tickets with the chance to win £50 in Amazon vouchers. After 5 min, participants were asked to stop playing.

Assessment of dependent variables. Participants were told they had to make several choices before their partner could start playing. They were reminded that their choices might improve or worsen their partner’s performance. Thus, if they selected stimuli that improved the partner’s performance they might lower their chances of receiving £50, whereas if the stimuli worsened the partner’s performance their likelihood of getting the prize would be higher. Before making their choices, participants were presented with different descriptions of the video game and different music clips that targeted specific emotions. For each description and for each music clip participants had to rate the extent to which they wanted their partner to read the description and listen to the clip before or while playing the game (*preferences for emotion-inducing stimuli*). Then, participants had to rate to what extent they wanted their partner to feel angry, fearful, or neutral (*explicit emotional preferences*). We always asked participants to indicate their preferences for emotion-inducing stimuli before their explicit emotional preferences to avoid demand characteristics. Next, participants had to rate to what extent they thought anger, fear, or a neutral emotion would be beneficial to success in the game (*perceived utility of emotions*). Finally, participants were fully debriefed.

Materials

Manipulation check

The Positive and Negative Affect Schedule (Watson et al., 1988) is a 20-item questionnaire that assesses participants' positive ($\alpha = .85$) and negative affect ($\alpha = .84$) on a 5-point Likert scale of 1 (*not at all*) to 5 (*extremely*).

The 12-item version of the Empathic Response Questionnaire (Batson et al., 1987) was used to assess participants' levels of empathic concern for and personal distress on behalf of Player A on a 7-point Likert scale of 1 (*not at all*) to 7 (*extremely*). Empathic concern was calculated as the average of the terms *warmth*, *soft-hearted*, *tenderness*, *moved*, *compassionate*, and *sympathetic* ($\alpha = .88$). Personal distress was calculated as the average of the terms *upset*, *grief*, *sorrow*, *distressed*, *worried*, and *anxious* ($\alpha = .87$).

Dependent variables

Preferences for emotion-inducing stimuli for the partner. Participants listened to two anger-inducing (Refuse/Resist by Apocalyptica; Mars from The Planets), two neutral (Treefingers by Radiohead; First Thing by Four Tet), and two fear-inducing (The Bon Dam by Julyan D; Hand of Fate: theme from the movie *Signs*) music clips, all used and validated by Netzer et al. (2015). Participants were also presented with three short game descriptions designed to elicit an angry, fearful, or neutral emotional state (Netzer et al., 2015). The anger-inducing game description described the main character fighting enemies after they had destroyed the character's village. The fear-inducing game description described the main character surrounded by dangerous enemies who want to kill him or her. The neutral game description described the main character monitoring his or her surroundings. For each description and music clip, participants were required to rate how much they wanted their partner to read or listen to it on a scale of 1 (*not at all*) to 7 (*extremely*). Given that the correlation between stimuli was high ($r = .63$, $p = .001$, anger-inducing stimuli; $r = .57$, $p = .001$, fear-inducing stimuli; and $r = .54$, $p = .001$, neutral stimuli) and that we did not find

differences when entering type of stimuli as a within-subject variable in subsequent analyses (see supplementary online materials), we averaged the responses to the music clip and game description stimuli for each emotion thereby creating an emotion-inducing stimulus preference score for each emotion (anger, fear, and neutral).

Explicit emotional preferences. Participants rated how much they wanted their partner to feel neutral, angry, or fearful whilst playing the game on a scale of 1 (*not at all*) to 7 (*extremely*).

Perceived utility of emotions. Participants rated the extent to which feeling angry, neutral, or fearful would be helpful to success in the game on a scale of 1 (*not very helpful at all*) to 7 (*extremely helpful*).

Results

Manipulation check

A multivariate analysis of variance (MANOVA) with perspective taking (other-oriented, objective) and goal (confrontation, avoidance) as independent factors and empathic concern and personal distress as dependent variables revealed a significant effect of perspective taking, $F(1, 139) = 13.79, p = .001, \eta_p^2 = .09$. Pairwise comparisons showed that participants in the other-oriented perspective taking condition ($M = 4.47, SD = 0.15$) reported significantly higher empathic concern than those in the objective perspective taking condition ($M = 3.70, SD = 0.14$); $p = .001$. For personal distress, results of a MANOVA did not show a significant effect of the perspective taking manipulation, $F(1, 139) = 2.91, p = .09, \eta_p^2 = .02$ (see Table 1 for means).

Table 1

*Mean and Standard Deviations of Vicarious Emotions and Emotion Utility by**Experimental Condition*

Variable	Other-oriented perspective		Objective perspective	
	Confrontation goal	Avoidance goal	Confrontation goal	Avoidance goal
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Empathic concern	4.18 (1.21) ^a	4.75 (0.96) ^a	3.62 (1.42) ^b	3.77 (1.29) ^b
Personal distress	2.20 (1.30) ^a	2.33 (1.02) ^a	1.92 (1.05) ^a	1.99 (0.94) ^a
Anger utility	5.69 (1.39) ^a	2.66 (1.64) ^b	3.44 (1.59) ^b	3.74 (2.11) ^b
Fear utility	3.34 (1.92) ^b	6.09 (1.31) ^a	3.31 (1.95) ^b	3.09 (1.99) ^b
Neutral utility	3.71 (1.91) ^a	2.60 (1.33) ^b	3.42 (1.86) ^a	3.03 (1.12) ^a

Note. Rows with different superscripts indicate statistically significant differences at $p < .05$.

A MANOVA with the dependent variables positive and negative affect showed that perspective taking, goal, and their interaction were non-significant for positive [perspective taking: $F(1, 139) = .97, p = .33, \eta_p^2 = .007$; goal: $F(1, 139) = .18, p = .67, \eta_p^2 = .001$; Perspective Taking \times Goal: $F(1, 139) = 1.14, p = .29, \eta_p^2 = .008$] and negative affect [perspective taking: $F(1, 139) = .63, p = .43, \eta_p^2 = .005$; goal: $F(1, 139) = .48, p = .49, \eta_p^2 = .003$; Perspective Taking \times Goal: $F(1, 139) = .58, p = .45, \eta_p^2 = .004$].

Main analyses

Explicit emotional preference. Figure 1a shows explicit emotional preferences by condition.

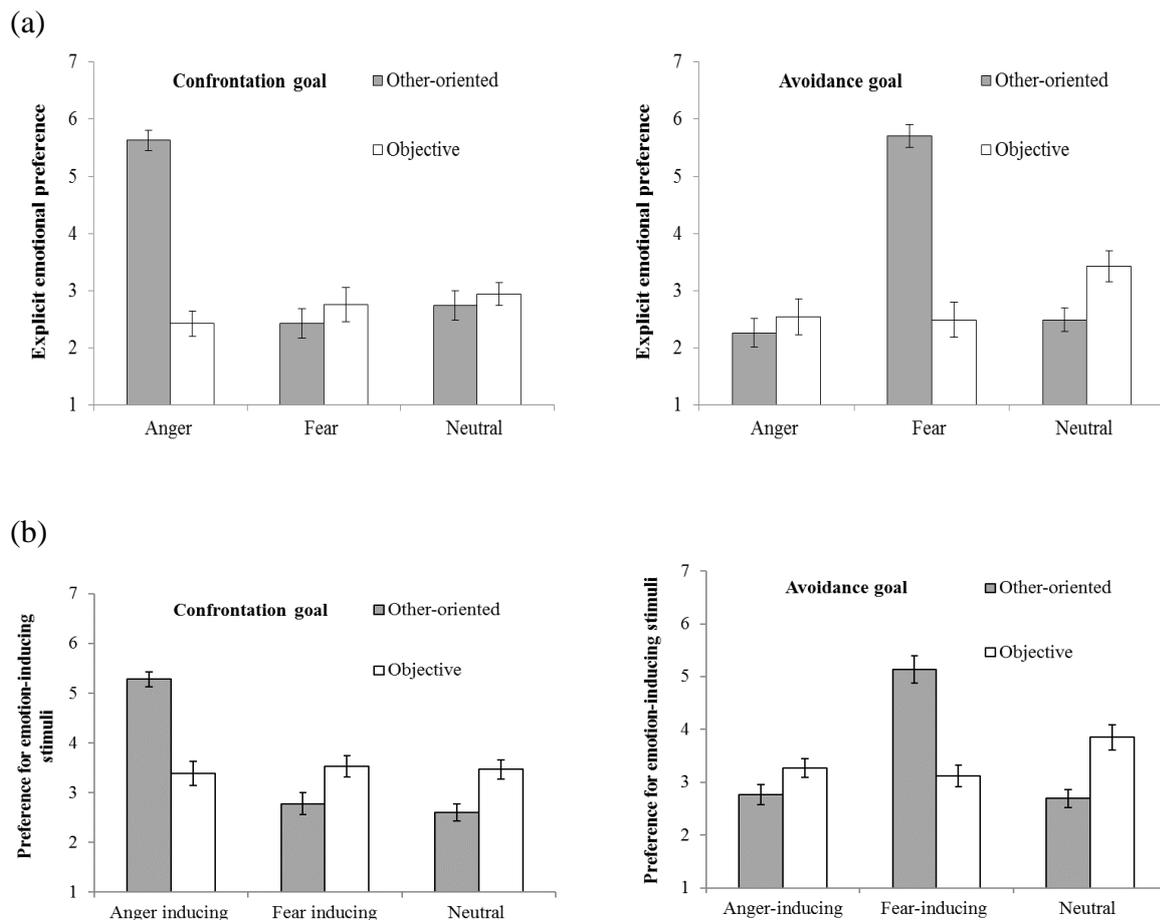


Figure 1. Mean of (a) explicit emotional preference and (b) preference for emotion-inducing stimuli. Error bars display standard errors.

A repeated-measures analysis of variance (ANOVA) with perspective taking (other-oriented, objective) and goal (confrontation, avoidance) as between-subjects variables and emotion (anger, fear, and neutral) as within-subject variable produced a significant Emotion \times Perspective Taking \times Goal interaction, $F(1, 136) = 54.13, p = .001, \eta_p^2 = .28$. In the confrontation goal condition, participants in the other-oriented perspective taking condition reported a higher preference for anger, $F(1, 69) = 125.17, p = .001, \eta_p^2 = .65$, than participants in the objective perspective taking condition. There were no differences between conditions for fear, $F(1, 69) = .65, p = .42, \eta_p^2 = .009$, and neutral, $F(1, 69) = .41, p = .53, \eta_p^2 = .006$. Pairwise comparisons showed that participants in the objective perspective taking condition did not differ in their ratings of explicit emotional preferences for the different emotions

($ps > .09$). However, participants in the other-oriented perspective taking condition reported a higher preference for anger compared to fear, $d = 3.20$, $SE = .30$, $p = .001$, and neutral, $d = 2.89$, $SE = .27$, $p = .001$. In the avoidance goal condition, participants in the other-oriented perspective taking condition reported a higher preference for fear, $F(1, 69) = 76.62$, $p = .001$, $\eta_p^2 = .53$, and less neutral, $F(1, 69) = 7.59$, $p = .01$, $\eta_p^2 = .10$, than participants in the objective perspective taking condition. There were no differences between conditions for anger, $F(1, 69) = .51$, $p = .48$, $\eta_p^2 = .007$. Pairwise comparisons showed that participants in the objective perspective taking condition did not differ in their ratings of explicit emotional preferences for the different emotions ($ps > .08$). However, participants in the other-oriented perspective taking condition reported a higher preference for fear compared to anger, $d = 3.46$, $SE = .31$, $p = .001$, and neutral, $d = 3.23$, $S.E. = .41$, $p = .001$.

Preferences for emotion-inducing stimuli for the partner. Figure 1b shows preferences for each emotion-inducing stimulus by experimental condition (descriptions and music clips combined; for separate analysis of music clips and game descriptions, please see supplementary material). A repeated-measures ANOVA with perspective taking (other-oriented, objective) and goal (confrontation, avoidance) as between-subjects variables and emotion (anger, fear, and neutral) as within-subject variable produced a significant Emotion \times Perspective Taking \times Goal interaction, $F(1, 136) = 35.45$, $p = .001$, $\eta_p^2 = .21$. In the confrontation goal condition, participants in the other-oriented perspective taking condition chose stimuli that were more anger inducing, $F(1, 69) = 43.88$, $p = .001$, $\eta_p^2 = .39$, less fear-inducing, $F(1, 69) = 4.42$, $p = .04$, $\eta_p^2 = .06$, and less neutral, $F(1, 69) = 11.70$, $p = .001$, $\eta_p^2 = .15$, than participants in the objective perspective taking condition. In the avoidance goal condition, participants in the other-oriented perspective taking condition chose stimuli that were more fear inducing, $F(1, 69) = 35.38$, $p = .001$, $\eta_p^2 = .34$ and less neutral, $F(1, 69) = 15.34$, $p = .001$, $\eta_p^2 = .18$ than participants in the objective perspective taking condition.

There were no differences for anger-inducing stimuli, $F(1, 69) = 3.69, p = .06, \eta_p^2 = .05$ (Figure 1b).

Perceived utility of emotions. A repeated-measures ANOVA with perspective taking (other-oriented, objective) and goal (confrontation, avoidance) as between-subjects factors and emotion (anger, fear, and neutral) as within-subject factor revealed a significant Emotion \times Perspective Taking \times Goal interaction, $F(1, 136) = 32.04, p = .001, \eta_p^2 = .32$. In the confrontation goal condition, participants in the other-oriented perspective taking condition perceived anger to be significantly more useful than participants in the objective perspective taking condition, $F(1, 69) = 39.83, p = .001, \eta_p^2 = .37$. There were no differences between perspective taking conditions for fear, $F(1, 69) = 0.01, p = .94, \eta_p^2 = .001$, or neutral, $F(1, 69) = 0.44, p = .51, \eta_p^2 = .01$. Participants expected anger to be more effective than fear, $d = 2.34, SE = 0.39, p = .001$, or neutral, $d = 1.97, SE = 0.41, p = .001$ (Table 1). In the avoidance goal condition, participants in the other-oriented perspective taking condition perceived fear ($F(1, 69) = 55.35, p = .001, \eta_p^2 = .45$) to be significantly more useful and anger ($F(1, 69) = 5.78, p = .02, \eta_p^2 = .08$) to be significantly less useful than participants in the objective perspective taking condition. There were no differences between perspective taking conditions for neutral, $F(1, 69) = 2.11, p = .15, \eta_p^2 = .03$. Participants expected fear to be significantly more useful than anger, $d = 3.42, SE = 0.39, p = .001$ (Table 1).

Testing Mediation of Explicit Emotion Preferences and Emotion Utility Beliefs.

Given that participants differed in their perception of emotion utility depending on the perspective taking condition, one may argue that participants' preferences for emotion-inducing stimuli may be driven by different emotion utility beliefs rather than explicit emotion preferences. In other words, rather than by altruistic affect worsening the results may additionally be explained by theory of mind, as participants in the other-oriented perspective taking condition may actually be better at anticipating what emotion may be more beneficial

for the target depending on the goal. To differentiate between these two alternative hypotheses, we conducted two moderated mediation analyses (i.e., one for preference of anger-inducing stimuli, one for preference for fear-inducing stimuli; see supplemental material for the depiction of the models) using the software *Mplus 7* (Muthén & Muthén, 2012). Specifically, we investigated whether emotion utility and/or explicit emotion preferences were significant mediators of the relationship between perspective taking and preferences for emotion-inducing stimuli across different goals (i.e., avoidance and confrontation). To formally test the mediation hypotheses, we used a bias-corrected bootstrap approach (1000 bootstraps) to create a 95% confidence interval (CI) around the mediated path (*ab*). This method was used as it has more power to detect mediation effects (Hayes & Scharkow, 2013). If the 95% lower and upper CI limits did not include zero, we concluded that the mediated effect was different from zero.

As depicted in the supplementary material, we ran a path analysis in which perspective taking (0 = objective, 1 = other-oriented), goal (0 = avoidance, 1 = confrontation), and their interaction terms were our exogenous independent variables. Explicit anger preference and anger utility belief were the mediators, and preference for anger-inducing stimuli was the final outcome. The effect of the interaction terms on both explicit anger preference ($B = 3.49$, $S.E. = .49$, $p < .001$) and anger utility belief ($B = 3.33$, $S.E. = .58$, $p < .001$) was statistically significant. Accordingly, we probed the effects of perspective taking on explicit anger preference and anger utility belief across the two goal conditions. Simple slope analysis indicated that for the avoidance condition the unstandardized effect of perspective taking on explicit anger preference ($B = -.29$, $S.E. = .41$, $p = .49$) was not significant, but it was for anger utility belief ($B = -1.09$, $S.E. = .47$, $p = .02$). Furthermore, only the effect of explicit anger preference ($B = .26$, $S.E. = .07$, $p < .001$) but not anger utility belief ($B = .06$, $S.E. = .06$, $p = .31$) was significantly related to preference for

anger-inducing stimuli. Next, the mediation analysis indicated that neither explicit anger preference ($ab = -.07$, 95% CI $[-.37, .11]$) nor anger utility belief ($ab = -.06$, 95% CI $[-.29, .04]$) significantly mediated the effect of perspective taking on preference for anger-inducing stimuli.

For confrontation, simple slope analyses indicated that perspective taking significantly predicted explicit anger preference ($B = 3.21$, $S.E. = .28$, $p < .001$;) and anger utility belief ($B = 2.24$, $S.E. = .35$, $p < .001$). Furthermore, only explicit anger preference ($B = .26$, $S.E. = .07$, $p = .001$) but not anger utility belief ($B = .06$, $S.E. = .06$, $p = .35$) predicted preference for anger-inducing stimuli. The mediation analysis showed that explicit anger preference mediated the effect of perspective taking on preference for anger-inducing stimuli ($ab = .83$, 95% CI $[.36, 1.34]$). Following Kline (2011), we conducted a further sensitivity analysis by constraining to the mediating paths to test whether the fit of the model was significantly different from the unconstrained one. Results showed that although the comparative fit index was good (i.e., .97), the chi square of the constrained model showed a significant increase ($\chi^2 = 9.03$, $df = 2$, $p = .01$) thereby attesting to the implausibility of the constraints.

For preference for fear-inducing stimuli, we entered as independent variables perspective taking, goal, and the interaction of both terms. Explicit fear preference and fear utility belief were entered as mediators, and preference for fear-inducing stimuli was the final outcome. The effects of the interaction terms were statistically significant on explicit fear preference ($B = -3.55$, $S.E. = .55$, $p < .001$) and fear utility belief ($B = -2.96$, $S.E. = .62$, $p < .001$). Accordingly, we probed the effects of perspective taking on explicit fear preference and fear utility belief across the two goal conditions. Simple slope analysis indicated that for the confrontation condition the unstandardized effect of perspective taking on explicit fear preference ($B = -.32$, $S.E. = .39$, $p = .41$) and on fear utility belief ($B = .04$, $S.E. = .45$, p

= .93) were not significant. In addition, neither explicit fear preference ($B = .19$, $S.E. = .10$, $p = .06$) nor fear utility belief ($B = -.08$, $S.E. = .09$, $p = .41$) were significantly related to preference for fear-inducing stimuli. The mediation analysis indicated that neither explicit fear preference ($ab = -.06$, 95% CI [-.34, .07]) nor fear utility belief ($ab = -.003$, 95% CI [-.15, .10]) significantly mediated the effect of perspective taking on preference for fear-inducing stimuli.

For avoidance, results showed that perspective taking significantly predicted explicit fear preference ($B = 3.23$, $S.E. = .38$, $p < .001$) and fear utility belief ($B = 3.00$, $S.E. = .40$, $p < .001$). As for confrontation, neither explicit fear preference ($B = .19$, $S.E. = .10$, $p = .06$) nor fear utility belief ($B = -.08$, $S.E. = .09$, $p = .41$) predicted a preference for fear-inducing stimuli. The mediation analysis showed that explicit fear preference mediated the effect of perspective taking on preference for fear-inducing stimuli ($ab = .62$, 95% CI [.004, 1.28]), but not fear utility belief ($ab = -.23$, 95% CI [-.88, .31]). Following Kline (2011), we applied an equality constrain to the mediating paths to test whether the fit of the model was significantly worse. Results showed that the model did not have a worse fit, as the χ^2 statistic was not significant ($\chi^2 = 3.13$, $df = 2$, $p = .21$). In the constrained model, the mediated path from perspective taking to preference for fear-inducing stimuli via explicit fear preference no longer significantly mediated the effect ($ab = .62$, 95% CI [.004, 1.28]). Likewise, fear utility belief was not a significant mediator ($ab = .15$, 95% CI [-.10, .42]).

Taken, together, these findings suggest that in confrontation only explicit anger preference was a significant mediator. For fear, we did not find a significant effect of our mediators. Therefore, our data did not provide empirical support for the alternative account, that is, the results cannot be explained by participants having a better theory of mind.

Discussion

Previous research has suggested that people may avoid making partners feel bad because such negative emotions are perceived as harmful (hedonic approach; Zaki & Williams, 2013). Alternatively, people may engage in interpersonal affect worsening to obtain a personal benefit (instrumental interpersonal affect worsening; Netzer et al., 2015). We showed that people may choose to be cruel to be kind. That is, agents may make a target feel bad to achieve a desired goal without the agents reaping any benefits themselves (altruistic affect worsening). Participants in the other-oriented perspective taking condition selected negative emotion-inducing stimuli that could benefit the target's performance in a video game, supporting the empathy hypothesis. Furthermore, participants' selection of stimuli was not random, as they wanted their partner to feel a specific negative emotion. One could argue that this is affect maintenance rather than worsening, given that the ostensible partner was initially described as being upset. However, affect maintenance seems an unlikely explanation. Although anger, fear, and upset are considered negative-valence emotions, they still differ in their levels of arousal (e.g., Feldman-Barrett, 2011).

Finally, participants perceived a particular negative stimulus to be more beneficial for succeeding in a particular game (supporting the beneficial goal hypothesis). Although one may argue that between-group differences in emotion utility belief indicated that participants in the other-oriented perspective taking condition were better at anticipating which emotion was more suitable for each goal (confrontation vs. avoidance), this alternative hypothesis was not supported. In two moderated mediation analyses only explicit emotion preferences were a significant mediator. Thus, participants were indeed cruel to be kind as they wanted the targets to experience a specific negative emotional response depending on the goal.

Our results support the extensive research on altruism and empathy showing how people help others even when altruistic behaviours may not entail a personal benefit (Batson

et al., 1988). This study extends these findings by revealing, for the first time, that people who experience empathic concern not only employ behaviour to decrease others' suffering but may also altruistically manipulate others' negative affect if this increases the others' long-term well-being, supporting the altruistic motivation hypothesis. Participants' efforts at worsening others' affect may additionally be motivated by egoistic reasons such as looking for social recognition (Cialdini et al., 1987) or demand effects. However, these explanations are unlikely to explain the findings of the current study, as participants were made aware that their choices were completely anonymous. Moreover, if participants' choices were driven by demand effects (i.e., participants selected stimuli similar to those they received), then similar patterns should have been obtained for participants in both perspective taking conditions. Finally, participants could have been motivated by restoring their own well-being (Hareli & Hess, 2010). While participants intentionally reduced their chances of receiving £50 by worsening their partner's affect, this may constitute only a low-cost action. Future research should therefore consider a higher cost, such as volunteering time (e.g., Batson, 2011).

The findings of our study pose a challenging question: What are the limits in affect worsening if it is for the sake of another's well-being? It may be that an agent will initiate the affect worsening process for another's well-being even if it is not necessary and even if the agent misperceives the other's need to feel bad to achieve long-term well-being (Hareli & Hess, 2010). Investigating the boundary conditions for altruistically and egoistically motivated interpersonal affect worsening will provide more information about adults' emotion–outcome expectancies regarding others' emotions, the cost–benefit calculations, and the factors they may consider when inducing a negative emotion in another. Below, we suggest a number of possibilities that can be explored in future research to test the boundary conditions of altruistic affect worsening.

Making sure that participants were unknown to each other was an intentional design choice of this study to maximize internal validity. Yet, research has shown that people may be more motivated to change others' affect in close relationships (Butler, 2011). To maximize external validity, future research could employ diary studies where people have to note information about the situation, the strategy, and the agent involved in the regulatory process (Parkinson & Simons, 2009). The use of video recording of dyads discussing real-life concerns or worries may also be helpful (Parkinson, Simons, & Niven, 2016). This line of research and the assessment of whether agents feel mixed emotions when engaging in affect worsening would clarify whether agents experience difficulties when worsening a target's affect depending on how close they feel to the target.

Other variables may impact agents' responsiveness to a target's emotions and goals. One of these variables may be empathic accuracy, the ability to optimally infer another's internal states (Ickes, 1997). High levels of empathic concern may lead to high responsiveness to a target only when empathic accuracy is high (Winczewski, Bowen, & Collins, 2016). Another variable that may impact agents' responsiveness is the perceptions of the target's regulatory skills (Parkinson & Simons, 2012). In fact, the experimental procedure used in the present study depicted the ostensible partner as upset and without any hope of getting over a breakup in order to provoke an empathic emotional reaction in the participants, which may have affected the participants' willingness to engage in affect worsening. Thus, future research could manipulate the agent's perception of the target's regulatory skills.

In this study, participants did not have the option to induce positive emotions in the target. Thus, we were unable to test if participants who experienced higher empathic concern might have wanted to increase the other's well-being by selecting positive emotion-inducing stimuli, as suggested by the empathy-altruism hypothesis (Batson, 2011). Future research may need to include happiness-inducing emotional stimuli to test this alternative hypothesis.

Furthermore, in our study and in the one by Netzer et al. (2015), participants could only worsen others' affect by selecting or modifying the situation. However, other research has investigated other strategies that can be used to change others' emotional states (e.g., co-rumination; Parkinson & Simons, 2012), and future research may benefit by studying these strategies. In sum, the present research opens an exciting research program that may enhance our knowledge of social cognition and interpersonal emotion regulation.

References

- Barraza, J. A., & Zak P. J. (2009). Empathy toward strangers triggers oxytocin release and subsequent generosity. *Annals of the New York Academy of Sciences, 1167*, 182–189.
- Batson, C. D. (2011). *Altruism in humans*. New York, NY: Oxford University Press.
- Batson, C. D., Duncan, B., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology, 40*, 290–302.
- Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., & Griffitt, C. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 55*, 52–77.
- Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin, 23*, 751–758.
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55*, 19–39.
- Batson, C. D., Kennedy, C. L., Nord, L. A., Stocks, E. L., Fleming, D. A., Marzette, C. M., ... Zenger, T. (2007). Anger at unfairness: Is it moral outrage? *European Journal of Social Psychology, 37*, 1272–1285.
- Butler, E. A. (2011). Temporal interpersonal emotion systems: The “TIES” that form relationships. *Personality and Social Psychology Review, 15*, 367–393.
- Carver, C. S. (2001). Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review, 5*, 345–356.

- 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*, 749–758.
- Erber, R., & Erber, M. W. (2000). The self-regulation of moods: Second thoughts on the importance of happiness in everyday life. *Psychological Inquiry*, *11*, 142–148.
- Feldman-Barrett, L. (2011). Constructing emotion. *Psychological Topics*, *3*, 359–380.
- Frijda, N. H. (1986). *The emotions*. Cambridge, England: Cambridge University Press.
- Gross, J. J. (2007). *Handbook of emotion regulation*. New York, NY: Guilford Press.
- Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3–24). New York, NY: Guilford Press.
- Hareli, S., & Hess, U. (2010). What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition & Emotion*, *24*, 128–140.
- Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science*, *24*, 1918–1927.
- Ickes, W. (1997). *Empathic accuracy*. New York, NY: Guilford Press
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus statistical modeling software (version 7)*. Los Angeles: Muthén & Muthén
- Netzer, L., Van Kleef, G. A., & Tamir, M. (2015). Interpersonal instrumental emotion regulation. *Journal of Experimental Social Psychology*, *58*, 124–135.
- Niven, K., Totterdell, P., & Holman, D. (2009). A classification of controlled interpersonal affect regulation strategies. *Emotion*, *9*, 98–509.

- Ohman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information processing mechanism. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 511–536). New York, NY: Guilford Press.
- Parkinson, B., & Simons, G. (2009). Affecting others: Social appraisal and emotion contagion in everyday decision making. *Personality & Social Psychology Bulletin, 35*, 1071–1084.
- Parkinson, B., & Simons, G. (2012). Worry spreads: Interpersonal transfer of problem-related anxiety. *Cognition and Emotion, 26*, 462–479.
- Parkinson, B., Simons, G., & Niven, K. (2016). Sharing concerns: Interpersonal worry regulation in romantic couples. *Emotion, 16*, 449–458.
- Tamir, M. (2009). What do people want to feel and why? Pleasure and utility in emotion regulation. *Current Directions in Psychological Science, 18*, 101–105.
- Tamir, M. (2015). Why do people regulate their emotions? A taxonomy of motives in emotion regulation. *Personality and Social Psychology Review, 1*–24.
- Tamir, M., & Ford, B. Q. (2009). Choosing to be afraid: Preferences for fear as a function of goal pursuit. *Emotion, 9*, 488–497.
- Tamir, M., Mitchell, C., & Gross, J. J. (2008). Hedonic and instrumental motives in anger regulation. *Psychological Science, 19*, 324–328.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.
- Winczewski, L., Bowen, J., & Collins, N. L. (2016). Is empathic accuracy enough to facilitate responsive behavior in dyadic interaction: Distinguishing ability from

motivation. *Psychological Science*. Advance online publication. doi:

10.1177/0956797615624491.

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion, 13*, 803–810.