On Drowning Children and Donations: Self-affirmation, Mood, and the Drowning Child

Argument – Exploring How to Increase Motivation to Donate

Fabienne Sandkühler

i6134822

MSc Human Decision Science

School of Business and Economics and Faculty of Psychology and Neuroscience

Primary Supervisor: Prof. Suzanne van Gils, Faculty of Psychology and Neuroscience

Secondary Supervisor: Prof. Elias Tsakas, School of Business and Economics

Maastricht University

2017

Word count: 12732

Table of Contents

Acknowledgements.....	2
Abstract.....	3
Introduction.....	4
Methods.....	15
Results.....	21
Discussion.....	34
References.....	47
Appendix.....	54

Acknowledgements

I would like to thank my primary supervisor, Prof. Suzanne van Gils, for her helpful feedback, guidance and answers to my many questions as well as her flexibility concerning my remote working and meeting over Skype. I would also like to thank my secondary supervisor, Prof. Elias Tsakas, for his helpful feedback. I thank both of them, my other tutors and Maastricht University for the informational and inspiring master's studies that I had the privilege to engage in. I would also like to express my gratitude to Lucius Caviola for sharing his preliminary results with me that are part of his PhD at Oxford University. Finally, I am thankful to my participants for taking the time to answer the survey.

Abstract

The aim of this study was to explore the effects and interactions of mood, self-affirmation and a moral argument for donating called *Drowning Child Argument* (DCA) on attitudes concerning charitable donations. The design was a between subjects, randomised 2x2x2. The survey population consisted of 298 adults. This design made it possible to test, among other things, if self-affirmation would decrease defensiveness about the Drowning Child Argument and thus lead to more favourable attitudes on donating. It did. Both self-affirmed and not self-affirmed participants in the DCA condition agreed significantly more that it was wrong not to donate than did control participants, but self-affirmed DCA participants agreed more than not self-affirmed DCA participants. Self-affirmed DCA participants reported a higher intent to donate than not self-affirmed DCA participants and self-affirmed control participants. The findings of this study support the idea that defensiveness prevents people from accepting the conclusion of the Drowning Child Argument. Implications for practice and future research are discussed.

Introduction

16,000 children under the age of five die each day of the consequences of extreme poverty (United Nations Inter-agency Group for Child Mortality Estimation, 2015) but could be saved with the help from people in developed countries, for example through charitable donations with short-term and long-term benefits for these children and their societies. Therefore, it is a matter of life and death to study how people's motivation to donate can be increased.

The *Drowning Child Argument* (Singer, 1972) is a widely used thought experiment to encourage people to conclude that donating a considerable part of their money to save lives follows from their ethical values. While there is ample anecdotal evidence (Singer, 1972; Singer, 2015), to date no published study has tested whether this thought experiment has the intended effect and if it can be increased. Because the Drowning Child Argument states that people are in one domain not following their own values, which is a criticism, and because people tend to also more in general feel uncomfortable when they notice they are being inconsistent, the Drowning Child Argument may be a threat to the self to which people respond defensively (e.g., Festinger, 1957; Steele, 1988). The present study investigates whether the Drowning Child Argument has the intended effect and whether two interventions to reduce defensiveness can increase the effect of this thought experiment. The two interventions are self-affirmation and a mood manipulation. While there has been a lot of research on the effect of self-affirmation on defensiveness (for reviews, see McQueen & Klein, 2006; Sherman & Cohen, 2006; Steele, 1988), there has been little research on the effect of self-affirmation on openness to moral arguments in particular. This study hopes to contribute to closing these gaps in the current literature and offer practical implications.

The Drowning Child Argument

The Drowning Child Argument is a thought experiment by philosopher Peter Singer (1972) that goes as follows: You see a small child drowning in a shallow pond. You could walk into the pond and save the child's life, but doing so would ruin your \$500 suit. You are faced with the decision of whether you want to save the child's life at the expense of \$500 or keep the money to yourself and let the child die. Singer then asks if it would be wrong not to save the child. When he asked his students or other audiences, the vast majority said that they would save the child and that it would be wrong to do nothing (Singer, 1972; Singer, 2010). He then asks if it would still be wrong to do nothing if the child was far away, or if she had a different nationality from yourself, or if there was a small chance that you would not succeed in saving her, or if there were people standing around the pond who could save her but do not do so. Still, the vast majority of people in the audiences he asked said that they consider it wrong to do nothing in these cases (Singer, 1972; Singer, 2010).

The Drowning Child Argument then proceeds to call people's attention to the fact that we can save a life, many lives in fact, by donating to effective charities. This encourages people to connect their existing values with the opportunity to help through donations. The Drowning Child Argument is a widely used thought experiment and many people have reported to me anecdotally that it helped them conclude that to be consistent with their own values, they want to spend less money on themselves and more money on those in need.

Caviola at the University of Oxford (manuscript in preparation) randomly assigned half of his participants to the full Drowning Child Argument, which included the main drowning child scenario (the story that you see a small child drowning in a pond nearby and you could ruin your suit and save her), three drowning child scenarios with one aspect changed in each (distance, bystanders, lack of full certainty) and a rationale for the

conclusion that it is similarly wrong not to donate than not to save the drowning child. For each scenario, participants were asked if they considered it wrong to do nothing, so that they would reflect upon whether they considered the factors distance, bystanders and full certainty to be morally relevant. As expected, participants agreed that it was wrong to do nothing in each scenario, which is the premise of the Drowning Child Argument and necessary for it to be able to exert the intended influence on donation attitudes. The other half of the participants received only the main story of the child drowning nearby and the rationale for the conclusion that it is similarly wrong not to donate than not to save the drowning child. They did not receive the drowning child scenarios in which an aspect was changed, nor were they asked if it was wrong to do nothing. Caviola found that participants who had been presented with the full thought experiment had significantly more favourable attitudes towards donating, for example they were more likely to think not donating was wrong than those who had not been presented with the full thought experiment. In the present study, one group of participants will be given the full Drowning Child Argument, while another group will receive a control story. Based on plausibility, Caviola's study (manuscript in preparation) and widespread anecdotal evidence, I hypothesise:

H1: Participants in the Drowning Child Argument condition will have more favourable attitudes towards donating than will participants in the control condition.

To further strengthen the Drowning Child Argument used in Caviola's study (manuscript in preparation), the Drowning Child Argument in the present study includes more and referenced information on the impact of effective donations and an additional drowning child scenario that I call the *psychophysical numbing scenario*. The additional information on the impact of donations is meant to educate participants about available

opportunities to donate in ways which, they can be confident, save and improve lives. This is to help participants translate the drowning child scenario in which there is a *small chance* that they will not succeed in saving the child to real world donations.

The psychophysical numbing scenario that is added to the other drowning child scenarios in the present study asks participants if they think it is wrong to do nothing, and to not save the child, if there is not just one child in the pond, but one hundred – and you are only able to save one. This addition is based on research by Fetherstonehaugh, Slovic, Johnson and Friedrich, 1997 and others (e.g., Friedrich et al., 1999; Slovic, 2007), showing that people value saving a life less as the number of lives at risk increase. In the psychological literature, this is known as *psychophysical numbing*. In other words, not only the absolute number, but also the proportion of lives saved out of lives at risk influences people's motivation to help. However, I hypothesise that when they are encouraged to reflect on this factor more explicitly, some people might conclude that they in fact think the proportion is not a very relevant moral factor. If psychophysical numbing is an obstacle to accepting the Drowning Child Argument that can be removed, it would be useful to include it also in practice. All participants receive the full Drowning Child Argument and are asked to rate each of the Drowning Child Argument scenarios, allowing tentative conclusions on the value of the psychophysical numbing scenario. Based on the strength of evidence for the influence of psychophysical numbing in decisions concerning lives at risk, I hypothesise:

H2: The psychophysical numbing scenario of the Drowning Child Argument will be rated as no less influential in changing one's mind about donating than the average of the other Drowning Child Argument scenarios.

Self-Affirmation

Self-relevant, critical information has benefits and costs. In the long-term, this information can allow us to better pursue our core values and goals (Dunning, 1995; Trope, 1975; Trope & Neter, 1994) but in the short-term, it can make us feel inadequate and insecure (Brown, 1990; Crocker & Major, 1989; Dunning, 1995; Steele, 1988).

Self-affirmation theory (Sherman and Cohen, 2006; Steele, 1988) posits that 1) people are motivated to protect their perceived integrity, 2) this can result in defensive behaviour and 3) perceived self-integrity is an overall self-concept that allows insecurities in one aspect of the self to be compensated by affirming another aspect of the self. Self-integrity has been defined as being a good and appropriate person overall (Sherman & Cohen, 2006). What is perceived as good and appropriate may differ across cultures, groups and situations (e.g., Quinn, 2006) and can involve characteristics such as intelligent, rational, autonomous, moral, being a good parent or a good friend. Self-affirmation refers to the assertion of the existence and value of one's self (e.g., Sherman & Cohen, 2006), for example, by writing an essay on a value important to you. According to self-affirmation theory, people can compensate perceived shortcomings in one domain by affirming success in another to maintain integrity. While self-affirmation theory suggests that self-affirmation reduces defensiveness through boosting positive self-concepts (Crocker, Niiya, & Mischkowski, 2008) and Tesser (2000) has argued that self-esteem maintenance processes are essentially interchangeable with self-affirmation processes, the evidence on this mechanism has been mixed (McQueen & Klein, 2006; Sherman & Cohen, 2006). Other researchers have instead found self-affirmation to be mediated by implicit positive affect (Koole et al., 1999), message scrutiny (Correll et al., 2004) and self-transcendence through love (Crocker, Niiya, & Mischkowski, 2008) as the mechanism through which self-affirmation reduces defensiveness.

Despite the debate on the mechanism through which it works, self-affirmation has been consistently found to decrease people's defensiveness (for reviews, see McQueen & Klein, 2006; Sherman & Cohen, 2006; Steele, 1988). A meta-analysis on the effects of self-affirmation on health behaviour change (Epton, Harris, Kane, van Koningsbruggen, & Sheeran, 2015) found that self-affirmation reliably improved the acceptance of health behaviour messages, intentions to change, and subsequent behaviour. For example, in one study (Sherman, Nelson & Steele, 2000), high and low caffeine consuming women were given an article with (false) information saying that caffeine consumption increased the risk for breast cancer. Half of the participants were asked to write a self-affirming essay. High caffeine consumers who were not self-affirmed were less willing to accept the article's argument and reduce their caffeine consumption than affirmed high caffeine consumers and low caffeine consumers. In another study (Sherman, Nelson & Steele, 2000), participants watched an AIDS education video. Half of the participants had self-affirmed by writing an essay on a value important to them. Affirmed participants reported being at higher risk for AIDS and were more likely to buy condoms offered to them after watching the video than were not affirmed participants.

A small number of studies have looked at self-affirmation in the context of moral decision making. In 2011, Stone, Whitehead, Schmader, and Focella found that highly prejudiced participants were significantly more interested in meeting an Arab-American based on his MySpace profile when the profile included questions encouraging self-affirmation. In another study (Wang & Zhao, 2017), participants were presented with an organ donation message that ended with a call to register as an organ donor. Self-affirmed participants showed significantly more favourable attitudes towards organ donation and a significantly greater intent to seek further information on registering as an organ donor.

Self-affirmed participants in a study by Harvey & Oswald (2000) who watched a civil rights video showed increased support for Black programs, whereas their non-self-affirmed counterparts did not. It has been proposed (Comer & Sekerka, 2017) that people should, among other things, self-affirm to help themselves sustain their moral courage at their workplace. However, this recommendation was based only on the general literature on self-affirmation and not on evidence on self-affirmation and moral behaviour in particular. The fact that self-affirmation can increase openness for a prejudice-challenging meeting, information on health risks, organ donation messages and interventions against racial discrimination, suggests that the present study might find self-affirmation to increase openness for a moral argument for charitable donating.

The present study includes a commonly used self-affirmation manipulation (Sherman & Cohen, 2006) in which half of the participants self-affirm by writing a short essay on the value they had ranked as most important to them from a list of values. This list did not include values directly related to donating, such as “morality”, “helping others” or “generosity”. This was intentional, as self-affirmation related to the threat to self-integrity that is provoking the defensive behaviour has been found to backfire, increasing defensiveness and close-mindedness (Blanton et al., 1997; Ehrlich & Gramzow, 2015; Jacks & O’Brien, 2004, Sherman & Cohen, 2006). I expect the questions on participants’ attitudes on donations to be sufficiently self-challenging by themselves to benefit from this defensiveness-reducing intervention, just like, for example, the question about whether participants would like to meet the Arab-American person benefited from self-affirmation in the study by Stone et al. (2011). Most importantly, I expect the effect of self-affirmation to be particularly pronounced in participants who are presented with the Drowning Child Argument, because the DCA is particularly challenging to self-integrity and if one is open to,

and thus more likely to accept, the argument, this has strong implications for one's attitudes on donating. As mentioned above, self-affirmation has been found to decrease defensiveness to persuading messages, which in turn affected responses (e.g., Sherman, Nelson & Steele, 2000; Wang & Zhao, 2017). Based on the aforementioned literature on self-affirmation and the self-challenging nature of the Drowning Child Argument and the questions on attitudes towards donating, I hypothesise:

H3: Participants in the self-affirmation condition will have more favourable attitudes towards donating than will participants in the no self-affirmation condition.

H4: Self-affirmation and Drowning Child Argument (DCA) will interact so that self-affirmed DCA participants will have particularly favourable attitudes towards donating.

Mood and Defensiveness

Self-affirmation is not the only intervention to be expected on theoretical and empirical grounds to make people less defensive. The *mood-as-a-resource hypothesis* (Aspinwall, 1998; Trope & Neter, 1994) states that happy people are more likely than neutral mood or sad people to more openly process information that is negative and highly self-relevant. This is because people who are in a happy state are thought to have the psychological resource to cope with the negative information while people in a neutral or negative mood do not have as much of this resource. Mood is defined as a subcategory of affect, as is emotion (Halmburger, Baumert, & Schmitt, 2015). Mood is different from emotion, in that mood is a more diffuse affective state (e.g., Halmburger, Baumert, & Schmitt, 2015) that is not tied to the interpretation or evaluation of a particular event and not directed at a specific object (e.g., Reisenzein, 1994). In addition, it has been argued that mood

tends to last longer than emotion, which is thought to have a clear peak and then lessen relatively rapidly (e.g., Halmburger, Baumert, & Schmitt, 2015).

Consistent with the mood-as-a-resource hypothesis, happy mood has been found to decrease defensiveness in numerous studies (see Trope & Fishbach, 2000; Trope & Neter, 1994). For example, in a study by Trope & Neter (1994), participants were asked to recall positive or negative events from their past. Participants were subsequently given the opportunity to receive feedback on a social sensitivity test they had taken earlier. Participants were told which subscales they had done well in and in which ones they had not performed well and they could choose which subscales to receive feedback on. Participants who had been asked to recall positive events asked significantly more often for negative feedback than did participants who had been asked to recall negative events.

Trope & Neter (1994), in another study, gave low and high caffeine consumers information about caffeine and later asked them to write down what they remembered. Consistent with mood-as-a-resource, high caffeine consumers who had been asked to recall positive past events reported more negative pieces of information about caffeine than high caffeine consumers who had been asked to recall negative past events. In contrast, low caffeine consumers did not recall more negative information when they were in a more positive mood and recalled more negative information when they were in a negative mood than did high caffeine consumers. The authors argue that low caffeine consumers did not feel the need to be defensive about negative information about caffeine in the first place, as this information was of low personal relevance to them.

The present study includes a happy and a neutral mood manipulation to test the effect of mood on defensive behaviour concerning the Drowning Child Argument. People's own morality has been consistently found to be highly relevant to their self-perception

(Strohminger & Nichols, 2014), so based on the mood-as-a-resource hypothesis it can be expected that mood will function as a resource that happy participants can use to deal less defensively with the moral argument presented in this study. Thus, I hypothesise:

H5: Participants in the happy mood condition will have more favourable attitudes towards donating than will participants in the neutral mood condition.

H6: Mood and Drowning Child Argument (DCA) will interact so that happy DCA participants will have particularly favourable attitudes towards donating.

Mood and Morality

The mood manipulation in this study might not just affect moral reasoning through lowering defensiveness towards the Drowning Child Argument and towards the questions on participants' attitudes on donating, but also have an independent effect on moral reasoning and behaviour. It has been suggested that positive mood leads to increased concern for others and thus, to altruistic actions (Isen, 1987; Penner, Dovidio, Piliavin, & Schroeder, 2005). Consistent with this idea, Halmburger, Baumert, & Schmitt (2015) summarise the literature on mood and helping as follows: "Research on helping has provided strong evidence that positive and negative mood, in contrast to a neutral mood, increase the probability that normative standards and values will guide behavior". Compared to neutral mood participants, many studies have found that participants who had a happy mood induced show a stronger intent to help others (e.g., Isen & Levine, 1972; Gaudine & Thorne, 2001; North, Tarrant, & Hargreaves, 2004; Wellen, Hogg, & Terry, 1998). However, a study by Niesta Kayser, Greitemeyer, Fischer, and Frey (2010) on help giving and moral courage suggests that the effect of mood depends on the type of helping that is required. Moral courage has been defined as "the ability and willingness to act according to one's important values, even in the

face of opposition, disapproval, and the danger of ostracism” (Staub, 2003). Niesta Kayser et al. (2010) found that happy and negative mood (vs neutral mood) increased helping when the request did not demand moral courage but when it did, mood did not make a difference.

However, not all the evidence suggests that happy people are more moral. The only study comparing the effects of an induced happy mood to a neutral mood on the *Ultimatum Game* found no differences in how much the two groups cared about fairness (Harlé & Sanfey, 2007). Ultimatum Games are games in which one participant receives a sum of money and has to make another participant an offer of how to divide the money between them. The receiver can decide whether to accept or reject the offer. If the receiver accepts, both receive their share. If the receiver rejects, they both receive nothing. Happy participants did not differ from neutral mood participants in the way they divided and accepted offers. Other moods do seem to affect this game, sadness (Forgas & Tan, 2013) and anger (Andrade & Ariely, 2009) leading to fairer offers and increased rejection rates. In addition, experimental research on the effect of happy mood on the *Defining Issues Test* (DIT; Rest, 1975), which tests moral reasoning, has produced mixed results. Higher scores on the DIT indicate enhanced moral reasoning according to Kohlberg’s stages of moral development (Kohlberg, 1971). In some studies (Olejnik & LaRue, 1980), happy mood participants scored more highly than neutral mood or sad participants; in other studies (Zarinpoush, Cooper, & Moylan, 2000), sad participants scored more highly than neutral or happy participants. In a study by Blumenthal (2005) positive and negative mood did not affect overall moral reasoning scores in the DIT nor when evaluating the morality of someone else’s action in a video. Interestingly, in a study by Pagano and DeBono (2011), happy participants scored significantly higher on the test than did sad participants, but only for participants who reported being affected more by internal attitudes and values and little by external factors of

the situation (low self-monitors). However, it is not possible to determine from this finding if happy mood increased the moral reasoning scores of low self-monitors or if sad mood worsened their scores, or both. The present study measures attitudes towards donating, which can be described as help giving that does not involve moral courage. Therefore, the literature on mood and morality strengthens the prediction that happy mood participants will have more favourable attitudes towards donating than will participants in the neutral mood condition.

Methods

Participants

A total of 298 participants were recruited successfully, 283 of them on the human intelligence platform MTurk¹ (location set to US or UK) and 15 through the researcher's personal network. There were 120 female and 176 male participants and 2 participants who did not specify their sex. Out of the 298 participants, 220 currently lived in the US, 11 lived in Germany, 1 lived in Austria, 1 lived in the Netherlands, 1 lived in France, and 64 lived in the UK. Participants were 34 years on average ($SD = 10,8$) and apart from 47 participants, they had completed or were enrolled in college education.

Participants were required to confirm that they were at least 18 years of age, not familiar with the philosophical thought experiment called the Drowning Child Argument and not very familiar or involved with the social movement called Effective Altruism. The latter requirement was a precaution to avoid participants of the researcher's personal network who were likely to have encountered reasoning very similar to that of the Drowning Child Argument and to avoid ceiling effects.

¹ Mechanical Turk, or MTurk for short, is an online marketplace where individuals or businesses can post paid tasks and individuals who sign up as workers can browse through tasks and complete them. Workers in the US comprise about half the MTurk worker population with demographics generally similar to the overall Internet population in the US (Iperiotis, 2009). MTurk workers in the US have a higher education level than the US average level (Iperiotis, 2010). MTurk is part of Amazon Web Services.

Procedure

Participants were given a brief description of the study and informed that it would take 20-25 minutes. Participants were informed that participation in the study was voluntary, that they could end their participation at any time without penalty, and that their responses were anonymous and would be used for research purposes only. Participants were asked to confirm that they would like to take part.

Participants were randomly assigned to the happy or neutral mood condition, to the self-affirmation or no self-affirmation condition and to the Drowning Child Argument or control condition and were presented with these conditions in this order.

Participants in the happy mood condition watched the first 2:30 minutes of the music video “Happy” by Pharrell Williams. Participants in the neutral mood condition watched the first 2:30 minutes of a video with no music showing small, neon-coloured fish swimming. The use of short film clips as a mood-induction technique has been successfully demonstrated in numerous studies (Gerrards-Hesse, Spies, & Hesse, 1994). Participants indicated if they had watched the video or if they could not play the video. For participants who reported not being able to play the video, the survey was ended prematurely and they were presented with a notice explaining that the survey was ended because the video was necessary for this experiment. Manipulation checks of mood were conducted (as described below, p. 18).

All participants were presented the following twelve values in random order and were asked to rank them in order of importance to them: *aesthetic appreciation, sense of humor, relations with family and friends, spontaneity, social skills, athletics, musical ability or appreciation, neatness or tidiness, physical attractiveness, creativity, managerial skills, and romantic values*. Participants in the self-affirmation condition were then asked to write about how their first ranked value was important to them and to explain how this value was

particularly important at a specific time in their life. Participants in the no self-affirmation condition were then asked to write about their ninth ranked value and to explain how this value might be important to the average person and how this value might be particularly important at a specific time in life. Manipulation checks of self-affirmation were conducted (as described below, p. 19).

Participants in the Drowning Child Argument condition received the following story: “You see a small child drowning in a shallow pond. You could walk into the pond and save the child’s life, but doing so would ruin your \$500 suit. You are faced with the decision of whether you want to save the child’s life at the expense of \$500 or keep the money to yourself and let the child die.” They were asked to what extent they agreed or disagreed that it would be wrong to do nothing. Then they were presented with five alterations of the original story and asked each time to what extent they still agreed or disagreed it was wrong to do nothing in these cases. The five alterations consisted of the four scenarios Caviola (q.v. page 11; paper in progress) used as well as one alteration in which there are a hundred children in the pond and you can only save one. Participants were also asked “Do you think that if you and your friends were born in a very poor country, receiving money from someone much better off could make a difference in your community? What might you use the money for?”. After that, they were presented with a text arguing that not donating is analogous to not saving the drowning child. This text also provided the referenced information that there are charitable organisations that have been scientifically proven to improve and save lives and gave two examples of the work of such organisations. A manipulation check of the moral argument was conducted (as described below, p.19).

Participants in the control condition were given a text of similar structure and length about a child playing happily alone. The control text makes it easier to rule out that any

differences between conditions were due to reading a text of a certain length, the structure of the questions and, importantly, priming with children. The full Drowning Child Argument condition and its control are given in Appendices A and B, respectively.

Following the Drowning Child condition or its control, participants were presented with the dependent variables of this study, consisting of three items on participants' attitudes towards donating, including perceived wrongness of not donating, intent to donate and perceived moral obligation to donate. Participants were asked to give their reasoning for their answer and, if they disagreed with the items, to say what piece of information, if true, would make them change their mind. Participants rated to what extent the questions concerning the wrongness of doing nothing in the Drowning Child main story and its alterations or the questions about the control story made them change their mind about donating when asked about it later.

Finally, participants provided demographic data, including their sex, age, nationality, current country of residence, monthly income, significant debts (yes/no), highest level of education achieved or in progress and if they have children. Participants were also asked if they knew the information about the impact of charities presented in the Drowning Child Argument condition of the study and to what extent they had thought about donating before this study. Participants were debriefed and thanked for their participation.

Measures

Mood manipulation checks. There were three manipulation checks for the mood manipulation – one simple question (“How happy do you feel right now?”), measured from 1 (*extremely unhappy*) to 7 (*extremely happy*) directly after watching the video and the same question and the I-PANAS-SF (International Positive and Negative Affect Schedule Short Form) at the end of the survey after measurement of the dependent variables. The

I-PANAS-SF consisted of ten questions asking to what extent participants felt a certain way, from 1 (*not at all*) to 5 (*extremely*). There were five questions measuring positive affect, namely, “How *active/alert/attentive/determined/inspired* do you feel right now?” and five questions measuring negative affect, namely, “How *afraid/ashamed/hostile/nervous/upset* do you feel right now?”. Two separate scores for positive and negative affect were computed.

Self-affirmation manipulation checks. There were two manipulation checks for the self-affirmation manipulation – one simple 5-point Likert-scale item from 1 (*strongly disagree*) to 5 (*strongly agree*) with the instruction “Please rate how satisfied you are with yourself in this moment”, measuring participants agreement with the statement “On the whole, I am satisfied with myself” directly after the self-affirmation manipulation and the 10-item Rosenberg self-esteem scale at the end of the survey after measurement of the dependent variables. The Rosenberg self-esteem scale was presented as a 5-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Five items were reverse coded and the average score of all items was computed. The full Rosenberg self-esteem scale can be found in Appendix C.

Moral argument manipulation check. The Drowning Child Argument can only work the way it is intended if participants agree that it would be wrong not to save the drowning child in the main story and its alterations. To get an impression of how well this criterion was met, the mean and standard deviation of agreement scores with this were computed. Participants reported their agreement to the following statements on a 7-point Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*):

1. “If we can easily save the life of an innocent person who is in danger at little cost to ourselves, it would be morally wrong for us to do nothing.”

2. “Even if there are other people who could help the person, but they do nothing, it would still be morally wrong for us to just watch and not help if we can.”
3. “Even if there is a small chance that we won't succeed in saving the life of the person in danger, it would still be morally wrong for us to do nothing.”
4. “Even if the person in danger is physically far away from us, it would still be morally wrong for us to do nothing if we could easily help.”
5. “Even if there were a hundred children drowning but we were only able to save one, it would still be morally wrong for us to do nothing.”

Dependent variables. The three dependent variables assessed participants' attitudes towards donating. All three dependent variables were measured on a 7-point Likert scale in which 1 indicates strong disagreement and 7 strong agreement. The first dependent variable, *wrongness of not donating*, stated: “How much do you agree with the following statement? ‘It is wrong NOT to donate some money to effective charities.’” The second dependent variable, *intent to donate*, stated: “If you do not have an income, imagine you would land a job average for your career path next month and that you are not in a particularly difficult financial situation. How much do you agree with the following statement? ‘Starting the month after next, I intend to donate a significant proportion, about 5 or 10%, of my income to whatever charities, projects or political campaigns I consider impactful.’” The third dependent variable, *moral obligation*, stated: “Do you think you are morally obligated to donate to help people in need?”

Rating of Drowning Child Argument items. Participants rated on a 5-point Likert scale from 1 (*not at all*) to 5 (*a great deal*) to what extent the questions concerning the wrongness of doing nothing in the Drowning Child main story and its alterations or the

questions about the control story made them change their mind about donating when asked about it later.

Statistical analyses

Statistical analyses were conducted using IBM SPSS Statistics 24. For each dependent variable, a 2x2x2 three-way ANOVA was conducted with mood, self-affirmation and Drowning Child Argument conditions as the independent variables. Simple main effects were computed for significant or trend interactions. A repeated measures ANOVA was used to test for differences between wrongness of not donating, intent to donate and moral obligation. A repeated measures ANOVA tested whether the rating of the Drowning Child Argument items differed significantly from one another and a one-sample *t* test compared specifically how influential the psychophysical numbing scenario was rated to be compared to the mean of the other four drowning child scenarios combined. An independent sample *t* test was used to compare the ratings of the Drowning Child Argument items to the rating of the items in the control condition. In addition, independent sample *t* tests were used to determine whether women and men differed significantly in their attitudes on donating. The objections participants who disagreed in the dependent variables raised are summarised and reported without further statistical analyses.

Results

Descriptive Statistics

A total of 298 participants completed the survey. The raw data set contained 357 responses but 57 of these were removed because participants did not complete the survey and 2 responses were removed because the participant's level of English was not sufficient. An

inspection of the written responses showed that the level of English of all other participants was not only sufficient but proficient.

Out of the 298 participants, 140 were in the happy mood condition and 158 in the neutral mood condition. There were 134 participants in the self-affirmation condition and 164 participants in the no self-affirmation condition. There were 147 participants in the Drowning Child Argument condition and 151 in the control condition. Table 1 shows how many participants were in each of the conditions in combination.

Table 1

Number of participants in each combination of conditions

mood condition	self-affirmation condition	Drowning Child Argument condition	N
happy mood	self-affirmed	DCA	35
		control	35
	not self-affirmed	DCA	31
		control	39
neutral mood	self-affirmed	DCA	32
		control	32
	not self-affirmed	DCA	49
		control	45
Total			298

Manipulation Checks

Mood manipulation check. The first two mood manipulation checks, that is to say, the question “How happy are you right now?” right after the manipulation and at the end of the survey were successful. The third mood manipulation check, the I-PANAS-SF at the end of the survey, was not successful. The null hypothesis of Levene’s test for equality of error variances for the second manipulation check was rejected, $p = .001$. Owing to this violated assumption, a t statistic not assuming homogeneity of variance was computed for that test.

The first manipulation check of the mood manipulation revealed that participants in the happy mood condition reported feeling significantly happier ($M = 5.31$, $SD = 1.17$) than participants in the neutral mood condition² ($M = 4.69$, $SD = 1.28$), $t(296) = 4.33$, $p < .001$, $d = 0.51$. A score of 4 represents the response *neither happy nor unhappy* and a score of 5 represents *slightly happy*. The second manipulation check, the same question but asked at the end of the survey, revealed that participants in the happy mood condition reported feeling significantly happier ($M = 5.00$, $SD = 1.22$) than participants in the neutral mood condition ($M = 4.63$, $SD = 1.44$), $t(296) = 2.39$, $p = .018$, $d = 0.28$.

The third mood manipulation check for the mood manipulation at the end of the survey, I-PANAS-SF, allowed a minimum score of 5 and a maximum score of 25 for positive and negative affect. An independent samples t test revealed that the I-PANAS-SF positive affect scores of participants in the happy mood condition ($M = 15.89$, $SD = 3.95$) and neutral mood condition ($M = 15.58$, $SD = 3.99$) did not differ significantly from each other³, $t(296) = 0.67$, $p = .501$. The negative affect scores of participants in the happy mood

² The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .088$.

³ The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .610$.

condition ($M = 6.86$, $SD = 3.07$) and the neutral mood condition ($M = 7.23$, $SD = 3.19$) did also not differ significantly⁴, $t(296) = 1.02$, $p = .310$.

Self-affirmation manipulation check. There were two manipulation checks for the self-affirmation manipulation. The first manipulation check, the item “On the whole, I am satisfied with myself” presented directly after the manipulation, was not successful. The second manipulation check at the end of the survey, the Rosenberg self-esteem scale which also included the item used as the first manipulation check, was also not successful.

The first manipulation check of the self-affirmation manipulation revealed that participants in the self-affirmation and the no self-affirmation conditions did not differ significantly in their responses to how satisfied they were with themselves⁵, $t(296) = 0.16$, $p = .875$. The second manipulation check of the self-affirmation manipulation revealed that participants in the self-affirmation and the no self-affirmation conditions did not differ significantly on the Rosenberg self-esteem scale⁶, $t(296) = 0.98$, $p = .327$.

Drowning Child Argument manipulation check. To get an impression of the extent to which participants agreed with the five statements regarding the drowning child in the Drowning Child Argument, I computed the average agreement score for each statement. Participants’ agreement with the five statements of the Drowning Child Argument is necessary for the thought experiment to be able to have the intended effect on attitudes on donating. The five statements were rated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree). On average, participants agreed or strongly agreed with all five drowning child statements. The average agreement scores for the statements concerning the drowning child were 6.62 ($SD = 1.04$), 6.64 ($SD = .87$), 6.55 ($SD = .96$), 6.24 ($SD = 1.18$),

⁴ The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .757$.

⁵ The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .426$.

⁶ The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .219$.

and 6.33 ($SD = 1.15$) out of 7, respectively. A score of 6 reflects the response *agree* and a score of 7 *strongly agree*.

Dependent Variables

The three dependent variables measured participants' attitudes towards donating. All three dependent variables were measured on a 7-point Likert scale in which 1 indicates strong disagreement and 7 strong agreement.

Wrongness of not donating. The first dependent variable assessed the extent to which participants agreed with the statement "It is wrong NOT to donate some money to effective charities". The hypothesis (H1) that participants in the Drowning Child Argument condition would agree significantly more strongly that it is wrong not to donate was confirmed. The hypothesis (H4) that there would be a significant interaction between Drowning Child Argument and self-affirmation, so that participants who were self-affirmed and presented with the DCA would agree particularly more strongly, was moderately supported by a trend in the expected direction. The hypotheses that participants in the happy mood condition (H5) and participants in the self-affirmation condition (H3) would agree significantly more that it was wrong not to donate and that mood and DCA would interact significantly in such a way that participants who were in both the happy mood and DCA conditions would agree particularly strongly (H6), were not supported. In addition, there was a trend for women to agree more strongly that it is wrong not to donate than did men.

A three-way ANOVA for wrongness of not donating with mood, self-affirmation and Drowning Child Argument as the independent variables revealed that⁷ the only statistically significant effect on wrongness of not donating was the main effect of the Drowning Child

⁷ The null hypothesis of Levene's test for equality of error variances for mood, self-affirmation and DCA for DV1 was not rejected, $p = .065$.

Argument (DCA), $F(1, 290) = 34.20$, $p < .001$, partial eta squared = .105. As hypothesised, participants in the DCA condition agreed more strongly ($M = 4.62$, $SD = 1.84$) with the statement that it is wrong not to donate than did participants in the control condition ($M = 3.40$, $SD = 1.70$), $d = 0.69$.

In addition, there was a strong trend for the interaction effect between self-affirmation and Drowning Child Argument (Figure 1), $F(1, 290) = 3.47$, $p = .064$, partial eta squared = .012. An analysis of simple effects showed that both self-affirmed participants, $F(1, 290) = 27.38$, $p < .001$, partial eta squared = .086 and not self-affirmed participants, $F(1, 290) = 8.70$, $p = .003$, partial eta squared = .029, exhibited a significant effect of the Drowning Child Argument (DCA) in the same direction (Figure 1). The effect of self-affirmation showed a trend in the expected direction for DCA participants, $F(1, 290) = 1.99$, $p = .160$, partial eta squared = .007. For control participants, the effect of self-affirmation was not in the expected direction and not significant, $F(1, 290) = 1.49$, $p = .223$. Consistent with the hypothesis (H4), the Drowning Child Argument had a stronger effect on self-affirmed participants than on not self-affirmed participants and the group with the highest score on this item was the self-affirmed + DCA condition.

Figure 1. Interaction effect for wrongness of not donating between self-affirmation and Drowning Child Argument (DCA) with $p = .064$. The y-axis indicates agreement from 1 (*strong disagreement*) to 7 (*strong agreement*). A value of 3 corresponds to the response *somewhat disagree* and a value of 5 corresponds to *somewhat agree*.

No further main effects or interaction effects of this three-way ANOVA reached statistical significance or showed a trend. An independent t test between female and male participants for wrongness of not donating revealed⁸ a trend for female participants to agree more strongly that it was wrong not to donate ($M = 4.24$, $SD = 1.87$) than did male participants ($M = 3.84$, $SD = 1.88$), $t(294) = 1.81$, $p = .071$, $d = 0.21$.

Intent to donate. The second dependent variable measured the extent to which participants agreed with the statement “Starting the month after next, I intend to donate a significant proportion, about 5 or 10%, of my income to whatever charities, projects or

⁸ The null hypothesis of Levene’s test for equality of error variances could not be rejected, $p = .891$.

political campaigns I consider impactful”. The hypothesis (H4) that there would be a significant interaction between Drowning Child Argument and self-affirmation, so that participants who were self-affirmed and presented with the DCA would agree particularly more strongly, was confirmed. The hypothesis (H3) that participants in the self-affirmation condition, whether they were presented with the DCA or not, would agree significantly more that it is wrong not to donate was not supported. Self-affirmation only had a significant effect on DCA and not on control participants and DCA only had a significant effect on self-affirmed and not on not self-affirmed participants. The hypotheses that participants in the happy mood condition (H5) and participants in the DCA condition (H1) would overall agree significantly more that it was wrong not to donate and that mood and DCA would interact significantly in such a way that participants who were in both the happy mood and DCA conditions would agree particularly strongly, were not supported. In addition, women showed a significantly stronger intent to donate than did men.

A three-way ANOVA for intent to donate with mood, self-affirmation and Drowning Child Argument condition as the independent variables revealed that⁹ the interaction effect between self-affirmation and Drowning Child Argument was significant, $F(1, 290) = 4.59$, $p = .033$, partial eta squared = .016. An analysis of simple effects showed that the effect of DCA was significant for self-affirmed participants, $F(1, 290) = 5.34$, $p = .021$, partial eta squared = .018, but not for participants who had not been self-affirmed (Figure 2), $F(1, 290) = 0.42$, $p = .517$. The effect of self-affirmation was significant for DCA participants, $F(1, 290) = 7.0$, $p = .009$, partial eta squared = .024, but not for control participants $F(1, 290) = 0.13$, $p = .717$. Thus, the hypothesis (H4) that self-affirmation and

⁹ The null hypothesis of Levene’s test for equality of error variances for mood, self-affirmation and DCA for DV2 could not be rejected, $p = .856$.

Drowning Child Argument would interact so that self-affirmed DCA participants would have particularly favourable attitudes towards donating was confirmed for intent to donate.

Figure 2. Interaction effect for intent to donate between self-affirmation and Drowning Child Argument (DCA) with $p = .033$. The y-axis indicates agreement from 1 (*strong disagreement*) to 7 (*strong agreement*). A value of 3 corresponds to the response *somewhat disagree* and a value of 5 corresponds to *somewhat agree*.

No other main effects or interaction effects of this three-way ANOVA were significant. However, the main effect of self-affirmation showed a trend in the expected direction. Participants in the self-affirmation condition ($M = 4.21$, $SD = 1.96$) showed a stronger intent to donate than participants in the no self-affirmation condition ($M = 3.84$, $SD = 1.89$), $F(1, 290) = 2.40$, $p = .103$, partial eta squared = .009. In addition, Participants in the DCA condition ($M = 4.12$, $SD = 1.97$) showed a stronger intent to donate than participants

in the control condition ($M = 3.89$, $SD = 1.89$), but this effect just missed the mark of p below .2 that I use to label differences as trends, $F(1, 290) = 1.60$, $p = .207$, partial eta squared = .005. The effect was also driven solely by the high intent to donate of self-affirmed DCA participants – as described above, DCA had no effect on intent to donate for not self-affirmed participants.

An independent t test between female and male participants for intent to donate revealed that¹⁰ female participants reported a significantly higher intent to donate ($M = 4.41$, $SD = 1.97$) than did male participants ($M = 3.73$, $SD = 1.86$), $t(294) = 2.99$, $p = .003$, $d = 0.35$.

Moral obligation. The third dependent variable measured the extent to which participants agreed with the statement “Do you think you are morally obligated to donate to help people in need”. The hypotheses that participants in the happy mood condition (H5), the self-affirmation condition (H3), and the Drowning Child Argument (DCA; H1) condition would agree more that they have a moral obligation to donate and that mood and DCA (H6) as well as self-affirmation and DCA (H4) would interact significantly were not supported. There was no trend for women and men to respond differently to this item.

A three-way ANOVA for moral obligation with mood, self-affirmation and Drowning Child Argument as the independent variables revealed that no main effect or interaction effect was significant nor showed a trend¹¹. An independent t test between female and male participants for moral obligation revealed that¹² there was no significant difference between the extent to which female and male participants reported having a moral obligation to donate, $t(294) = -.34$, $p = .734$.

¹⁰ The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .851$.

¹¹ The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .314$.

¹² The null hypothesis of Levene’s test for equality of error variances was not rejected, $p = .551$.

Comparison of the three dependent variables. Participants agreed significantly more with moral obligation to donate than with wrongness of not donating and intent to donate. The latter two did not differ significantly from each other. A repeated measures ANOVA revealed that there were significant differences between the three dependent variables perceived wrongness of not donating, intent to donate and perceived moral obligation., $F(1, 297) = 6.51, p = .002$. This statistic was Greenhouse-Geisser corrected, because the assumption of sphericity was violated, $\chi^2(2) = 12.25, p = .002$. Post-hoc paired-samples t test revealed that perceived wrongness of not donating and intent to donate did not differ significantly from each other, $t(297) < .001, p = 1.00$. Participants agreed significantly more with perceived moral obligation ($M = 4.37, SD = 1.87$) than they did with perceived wrongness of not donating ($M = 4.00, SD = 1.93$), $t(297) = 3.50, p = .001$, and intent to donate ($M = 4.00, SD = 1.87$), $t(297) = 2.98, p = .003$.

Rating of the Drowning Child Argument Items

Participants in the Drowning Child Argument condition rated to what extent the different Drowning Child Argument scenarios, the question on whether donated money would make a difference to them if they were born in a very poor country and what they would use donated money for, and the information on the impact of charities presented to them had changed their mind about donating when asked about it later. The ratings of the different Drowning Child Argument scenarios, the additional question and the information on the impact of charities did not differ significantly. The rating of the influence of the psychophysical numbing scenario did also not differ significantly from the mean of the other four drowning child scenarios combined, supporting hypothesis H2. The Drowning Child Argument items were all rated as significantly more influential than the items of the control condition.

A repeated measures ANOVA revealed that the ratings of these seven items did not differ significantly from one another, $F(1, 146) = 1.43, p = .200$. Mauchly's test showed that the assumption of sphericity was violated, $\chi^2(20) = 117.13, p < .001$, so the above ANOVA uses the Greenhouse-Geisser corrected statistics. A one-sample t test revealed that the ratings of how much the psychophysical numbing scenario made participants change their minds did not differ significantly from the average of the four drowning child scenarios combined, $t(146) = -1.03, p = .305$. Independent sample t tests revealed that the questions in the control condition ($M = 0.26, SD = .64$) were rated as significantly less influential than the average of the DCA items ($M = 1.73, SD = 1.29$), $t(296) = 12.45, p < .001$. Levene's test revealed that the null hypothesis of equality of error variances was rejected, $p < .001$, so the given t statistic does not assume equality of variances.

Reasons for Disagreement

For each of the three dependent variables, participants briefly explained their reasoning for their answer, and if they objected, what piece of information, if true, would be most likely to change their mind. For the wrongness of not donating item, the most common primary reason why participants disagreed that not donating is wrong was that this does not apply to people who cannot afford to donate. This was true for 5 out of the 33 people who disagreed strongly, 19 out of the 48 people who disagreed, and 17 out of the 42 people who somewhat disagreed.

The second most common primary reason participants who disagreed it was wrong not to donate gave can be summarised with the three statements "people can do what they want with the money they earn/their money belongs to them/it is their personal choice". This was given as the primary reason by 7 out of the 33 people who disagreed strongly, 6 out of the 48 people who disagreed, and 5 out of the 42 people who somewhat disagreed.

The third most common primary reason why participants disagreed it was wrong not to donate was because they thought there were no or almost no effective charities/ donations are not guaranteed to do good/donations only lead to dependency and people “need to learn to support themselves”. This was the primary reason for 8 out of the 33 people who disagreed strongly, 3 out of the 48 people who disagreed, and 3 out of the 42 people who somewhat disagreed.

A small number of participants disagreed because they thought the wrongness of not donating item was about one person giving money to every single existing effective charity or every single effective charity that asks you to (6 participants); because they thought it was not necessary to donate anymore if you do good in a different way (2 participants); or because they considered donating to be supererogatory, i.e. they thought it was morally good but not morally necessary (1 person).

Participants who disagreed it was wrong to donate for the most part answered the question “What piece of information, if true, would be most likely to change your mind?” consistently with their reported reasoning. So, for example, those who disagreed because the statement did not apply to people who cannot afford to donate, said that information on the financial situation of someone could change their mind so that they would think not donating would be wrong in that person’s case. Participants who disagreed because they thought charities were not effective said that information proving the effectiveness of charities would change their mind. A related answer was given by a number of participants with different primary reasons for disagreeing. They said that if their donation would save someone who would otherwise die, they would change their mind. Participants who disagreed because “people can do what they want with the money they earn”, tended to say that nothing would change their mind.

Most participants who disagreed with the second item, intent to donate, did so because they did not have enough money right now or would need to save money first or because they thought the 5-10% of one's income each month asked about in this item was too high. Accordingly, most disagreeers said that having more money would make them change their minds. The second most common information that participants reported would make them change their minds on the intent to donate item was information that made them confident about the impact they would have with their donation. The responses to the third dependent variable, moral obligation, were very similar to the wrongness of not donating item.

Discussion

Summary of the Findings

This study investigated the effects of mood, self-affirmation and a moral argument for donating called *Drowning Child Argument* on three attitudes on donating: perceived wrongness of not donating, intent to donate and perceived moral obligation to donate. The hypothesis most supported by the results of this study was that self-affirmation and the Drowning Child Argument would interact in such a way that participants who were both self-affirmed and had been presented with the Drowning Child Argument would show attitudes that particularly favoured donating. This interaction was significant and in the predicted direction for intent to donate, and showed a strong trend for perceived wrongness of not donating. The hypotheses that received the least support were those concerning mood. There were no significant or trend main effects of mood or interactions of mood and the moral argument for any of the three attitudes on donating. The hypothesis that the Drowning Child Argument would have a main effect on attitudes towards donating received moderate

support. It had a highly significant main effect on perceived wrongness of not donating, but no significant or trend main effects on intent to donate and perceived moral obligation. Contrary to predictions, self-affirmation did not affect participants who had not been presented with the Drowning Child Argument. In addition, an exploratory analysis revealed that women agreed more strongly that it is wrong not to donate and showed a higher intent to donate than did men. This is consistent with previous studies (Charities Aid Foundation, 2011; Piper & Schnepf, 2008).

Discussion of the Mood Manipulation

The manipulation of mood seems to have been successful: Participants who watched the happy video reported being highly significantly happier right after watching the video than participants who watched the neutral video. It is interesting to note here that the effect of the mood manipulation persisted, although to a decreased degree, at the end of the survey when measured with the simple question “How happy do you feel right now?” but the mood manipulation seemed to have no effect on the positive nor the negative affect score of the I-PANAS-SF, a questionnaire with five questions on positive affect and five questions on negative affect. There are several reasons that might explain this pattern.

One explanation is that the effect of the mood manipulation lasted until the end of the survey, but the words *active*, *alert*, *attentive*, *determined*, and *inspired* on the one hand and *afraid*, *ashamed*, *hostile*, *nervous*, and *upset* did not capture the effect the mood manipulation had. Alternatively, it might be that the effect of the mood manipulation had in fact disappeared by the end of the survey, but participants in the happy mood condition reported being happier on the “How happy do you feel right now?” question, because they had answered exactly the same question in the survey before and wanted to appear consistent. I favour the first explanation, because in other studies, participants did not seem to feel the

need to give the same answer to mood questions (e.g., Raghunathan & Trope, 2002) and because the choice of items in the I-PANAS-SF seems indeed somewhat questionable for measuring positive and negative affect, as it does not include the concepts happy and sad. This questionnaire was included to measure affect with a standardised questionnaire, but it might be that the simple question “How happy are you right now?” was more fitting for this study.

In contrast to the self-affirmation manipulation, the mood manipulation did not have an effect on participants’ attitudes on donating by itself nor in interaction with the Drowning Child Argument. This might seem surprising, because in previous studies, mood was successful in decreasing defensiveness in very similar situations as self-affirmation. However, looking at two particularly comparable measures of the effect of mood and self-affirmation manipulations sheds some light on why the mood manipulation did not work in the present study while the self-affirmation manipulation did. Raghunathan and Trope (2002) and Sherman, Nelson and Steele (2016) studied low and high caffeine consumers’ defensiveness concerning an article on the supposed health risks of caffeine. The study by Raghunathan and Trope (2002) included a happy and a neutral mood condition. The study by Sherman, Nelson and Steele (2016) included a self-affirmation and a no self-affirmation condition. Both studies included a measure of intent to decrease caffeine consumption and found the manipulation of self-affirmation and mood respectively to have a significant effect on this measure. These studies do not report their effect sizes directly, but Cohen’s d could be calculated using the given means and standard deviations in one case and the F statistic and degrees of freedom in the other case. For high caffeine consumers, the mood manipulation of Raghunathan and Trope (2002) had an effect of $d = 0.69$ on intent to decrease consumption while the self-affirmation of Sherman, Nelson and Steele (2016) had an effect of $d = 1.96$ on

intent to decrease consumption. This comparison suggests that self-affirmation manipulations might be more effective in reducing defensiveness than are mood manipulations. To get a more intuitive grasp of what this difference in effect size, assuming the two studies are comparable enough, means: If we chose a participant from the control group and one from the manipulation group at random, in the self-affirmation study the probability that the person from the manipulation group reported a higher intent to decrease caffeine consumption than the control group person would be 92%, while in the mood study this probability would be only 69%. What can we learn from this little excursion on the mood manipulation? Firstly, it is useful to mention the effect sizes of your findings. Secondly, the mood manipulation in the present study might have shown no effect on attitudes on donating because inducing a happy mood is generally less effective in making people more open to highly self-relevant negative information than is self-affirmation, which is consistent with previous literature. This could be because self-affirmation targets the self-concept more directly than does a mood manipulation, and it is this self-concept in particular and not happiness more in general that highly self-relevant negative information challenges. In addition, more time passed between the mood manipulation and the dependent variables than between the self-affirmation manipulation and the dependent variables, possibly further amplifying the difference in effectiveness between the two manipulations.

However, all of this considered, there is still room for a further explanation of why there was no effect of mood on donation attitudes in this study. An implication of this study for mood and morality, therefore, might be that happy mood does not only fail to have an effect on helping when the helping involves moral courage (Niesta Kayser et al., 2010), but also in the domain of substantial donations. Donating a substantial proportion of one's income does not fit the definition of moral courage, but they both require costly helping

behaviour and this might have been what led to the similar results for attitudes on donating in the present study and moral courage in the study by Niesta Kayser et al. (2010). This could be because people might be motivated to protect their good mood and high-cost helping does not serve to do this (Niesta Kayser et al., 2010).

Discussion of the Self-Affirmation Manipulation

The self-esteem manipulation checks of self-affirmation were not successful. There were no significant differences between the self-affirmation condition and the no self-affirmation condition for the item “On the whole, I am satisfied with myself” immediately following the self-affirmation essay, nor was there an effect of the manipulation on the Rosenberg self-esteem scale at the end of the survey. However, this does not necessarily mean that participants in the self-affirmation condition were not successfully self-affirmed. While some studies on self-affirmation found that the effect of self-affirmation was mediated by self-esteem (Cohen et al., 2000; Fein & Spencer, 1997) and Tesser (2000) has argued that self-esteem maintenance processes are essentially interchangeable with self-affirmation processes, other studies found that self-affirmation was successful in making people less defensive, but did not affect their self-esteem. Instead, the effects of self-affirmation on defensive behavior have been found to be sometimes mediated by implicit positive affect (Koole et al., 1999), message scrutiny (Correll et al., 2004) and self-transcendence (Crocker, Niiya, & Mischkowski, 2008).

For example, just like in the present study, Schmeichel and Martens (2005) self-affirmed participants by having them write about an important personal value. Self-affirmed participants showed less worldview defense and death-thought accessibility than not self-affirmed participants, but no difference in self-esteem. The concept of self-esteem seemed like the most likely candidate to be used in a manipulation check,

although it probably would have been better to have used a state self-esteem measure instead of a measure of overall self-esteem. Additionally including an array of different manipulation checks for self-affirmation and exploring them in more depth would have been interesting but would have brought its own difficulties. It would have made the study even longer, and if these measures had been placed towards the end of the survey, they might not have detected anything, because of the time between the manipulation and the checks, and because the Drowning Child Argument and the dependent variables might have absorbed the effect of the manipulation, as has been found to happen with mood (Raghunathan & Trope, 2002). If the manipulation checks had been placed not at the end of the survey but after the manipulation, this would have weakened the manipulation by making the time between the manipulation and the dependent variables longer, and it was already not short. As will be discussed in the following, the self-affirmation manipulation had an effect on defensive behaviour concerning the moral argument presented to participants, which supports the validity of the self-affirmation manipulation.

Half of the participants wrote a (very short) essay on the value they had listed as most important to them. This was intended to self-affirm them, i.e. affirm to themselves who they are and what is important to them, strengthening their perceived self-integrity. If one reason why people do not change their attitudes when presented with the Drowning Child Argument and information on the impact of donations is that in order to protect their perceived self-integrity, they dismiss the argument without processing it open-mindedly (see Sherman and Cohen, 2006; Steele, 1988), then we should expect that strengthening the self to lead to more attitude change when presented with the moral argument. In other words, if defensiveness is the problem, then making people feel less need to be defensive should lead to more change. This idea is supported by this study's finding that self-affirmation and the

Drowning Child Argument interacted so that participants who were self-affirmed and had been presented with the Drowning Child Argument showed particularly favourable attitudes towards donating. This was the case for perceived wrongness of not donating and intent to donate. This finding suggests that the use of self-affirmation interventions can be extended to reduce defensiveness to moral arguments, an area that self-affirmation research had not studied before.

I expected defensiveness to not only play a role in processing the Drowning Child Argument (DCA), but also in processing the questions on attitudes towards donating. That is to say, I expected the questions (e.g. whether it is wrong not to donate) alone to be perceived as somewhat self-threatening to some participants and elicit defensive behaviour that would be reduced by self-affirmation. This idea was not supported by the current findings. While there was a trend for self-affirmed participants overall to report a higher intent to donate than not self-affirmed participants, this trend was solely due to self-affirmed DCA participants reporting a much higher intent to donate than all other combinations of the self-affirmation and DCA conditions. Self-affirmation did not affect the intent to donate of control participants. Self-affirmation did also not have a significant or trend main effect on perceived wrongness of not donating or on perceived moral obligation to donate. Instead, self-affirmation only had an effect in combination with the Drowning Child Argument. This suggests that self-affirmation lowered defensiveness, but this only made a difference when presented with an argument persuading them – self-challenging questions on this topic were not enough.

Discussion of the Drowning Child Argument

The manipulation check of the Drowning Child Argument was successful. On average, participants (strongly) agreed with the premises of the moral argument. I expected

participants to not be a hundred percent defensive about the Drowning Child Argument (DCA), even if they were not in the self-affirmation condition. So, the moral argument should have had an impact not only on self-affirmed, but also on not self-affirmed participants. This idea was supported by a highly significant stronger agreement that it was wrong not to donate in the Drowning Child Argument condition overall as well as more specifically, by a significantly stronger agreement in self-affirmed and in not self-affirmed participants in the Drowning Child Argument condition compared to participants in the control condition. There was no effect of DCA on not self-affirmed participants for the other two dependent variables intent to donate and perceived moral obligation.

The ratings of the different items in the Drowning Child Argument condition (five scenarios, information on the impact of donations and an additional question on whether donations would be helpful to participants themselves if they had been born in a very poor country) did not differ significantly from each other. The influence of the psychophysical numbing scenario was not rated as significantly different from the influence of the average of the other four drowning child scenarios combined. The items in the DCA condition were rated as significantly more influential than the control condition items. Thus, it seems that participants were aware that the moral argument had an influence on their thinking. A tentative interpretation of the rating is that all Drowning Child Argument items were in fact similarly influential, including the psychophysical numbing question that was new in this study. The fact that the items in the DCA condition were rated higher than the items in the control condition indicates that the items were not rated thoughtlessly. However, it is surprising that the DCA items were all rated as equally influential. It seems plausible that participants were not aware of the exact influence the items had on their thinking or did not put enough time into analysing which items were most influential, because they were asked

to do the rating at the end of the relatively long survey. It would therefore be interesting for future research to give the different scenarios of the Drowning Child Argument and the information about the impact of donations to different groups of participants separately to see how influential they really are.

Comparison of the Three Dependent Variables

The results differed considerably across the three dependent variables, different hypotheses receiving or lacking support from different dependent variables. The third dependent variable, perceived moral obligation, seemed quite unresponsive to the manipulations and different from the first two dependent variables. There were no significant or trend main effects or interactions effects at all for perceived moral obligation. Participants agreed highly significantly more with the moral obligation item than they did with the first two items (wrongness of not donating, intent to donate). This seems surprising, as the term “moral obligation” can be thought of as stronger than wrongness of not donating and, one could argue, even than intent to donate. It might be that participants agreed more with the moral obligation item because it included the phrase “to donate to help those in need”. In particular, the words “in need” might have evoked emotions favouring agreement. This might also have dominated the responses, minimising the effect of the manipulations. However, these explanations are very speculative.

Strengths and limitations

The moral argument in the present study encouraged people to connect their values to their wallets. Effective donations make a great difference to the lives of many people, so this is an important topic to study. This is the first research testing the effect of self-affirmation and mood on openness to a moral argument. Being open-minded about moral arguments is

important, because this helps us do what we think is right, which is valuable to ourselves and anyone affected by our actions. This is also the first research comparing the effect of the Drowning Child Argument (DCA) to a control without any part of the DCA and the first research to explore the relative merits of the different components of the DCA, including a new scenario based on the psychophysical numbing literature. The present study was also able to extend Niesta Kayser et al.'s (2010) finding that positive mood does not affect moral courage to high-cost donations. Thus, the design of this study allowed its findings to contribute new insights to the self-affirmation, mood, and morality literature.

There are two main limitations to this study. The first is based on the qualitative assessment of the reasons participants gave for disagreeing that it was wrong not to donate (DV1). The most common primary reason participants disagreed was that they thought the statement should not apply to people who cannot afford to donate. This statement was not meant to include people who are poor themselves and in hindsight it seems clear that it would have been better to have made this more explicit. Similarly, the most common reason why people disagreed with the intent to donate item (DV2) was that they would need to save some money first before they could start donating. While the intent to donate item asked participants to imagine they had a job and were not in a difficult financial situation, apparently it would have been better to make it even more clear that they should assume they were financially safe. Making this very clear might be particularly important when participants live in the US (compared to many European countries), as most of the participants in the present study did. This is because the US population has a weaker social security net and people may need to save much more money than e.g. Germans do to make sure they are covered in the case of, for example, a medical emergency. In fact, several

participants explicitly mentioned saving for treatment when disagreeing with the intent to donate item.

Secondly, roughly half of participants in the self-affirmation condition did not write more than three sentences about their chosen value and how it was important in their life. While self-affirmation interacted with the moral argument as expected, the effect sizes could have been bigger. Based on the previous literature on self-affirmation, it seems likely that there would be bigger effect sizes for this interaction in people who self-affirm more extensively, for example by writing a longer essay that takes at least 10 minutes. In addition, ordering values according to their personal importance already constitutes a weak form of self-affirmation (Wang & Zhao, 2017), which decreases the difference between the self-affirmation conditions. As what we are after when attempting to decrease defensiveness to moral arguments is a substantial – and not just statistically significant – effect, it might have been better to have encouraged extensive (as opposed to brief) self-affirmation. It could be interesting for future research to test if more extensive self-affirmation can have a substantial interaction effect with moral arguments on altruistic motivation.

A further limitation of this study is that it only looked at *attitudes* towards donating, which are only the first step to action. Future research could investigate whether interventions against defensiveness in combination with moral arguments like the Drowning Child Argument actually lead people to donate more. In addition, even if people donated more at first, they might not necessarily continue to do so. For this, further sustaining factors might be necessary, such as fulfilling feedback about the impact their donations are having and, if the donations are to be a substantial proportion of one's income, a social network supporting this decision. In addition, ideally, this study should be replicated with those who are currently in a good position to donate – financially very comfortable people, who are not likely to be

recruited through MTurk. Finally, the participants in this study were primarily American, and results might differ in other cultures.

Conclusion

Overall, the hypotheses that being presented with the moral argument by itself would lead to more favourable attitudes towards donating and that self-affirmation would decrease participants' defensiveness towards the moral argument, leading to particularly favourable attitudes towards donating for participants who received this combination of interventions, received moderate support. The hypothesis that self-affirmation by itself and happy mood by itself would lead to more favourable attitudes towards donating and the hypothesis that mood in interaction with the moral argument would lead to favourable attitudes towards donating received no support.

What can we learn from these findings? The findings suggest that research on self-affirmation should be more interested in moral arguments, and that research on morality should be more interested in self-affirmation. This study was able to extend the finding by Niesta Kayser et al. (2010) that moral courage is not affected by positive mood to high-cost helping in the domain of donations. This has implications for practice, in that organisations and individuals who are interested in encouraging substantial donations (or also moral courage) cannot rely on positive mood to increase people's motivation but may instead self-affirm those they are trying to persuade.

Apart from including two defensiveness interventions (mood and self-affirmation), which were the most important extensions of Caviola's study (manuscript in progress), this study also extends the work of Caviola by including a control group who did not receive any part of the Drowning Child Argument. This makes it possible to assess the full impact of this thought experiment. Caviola's study was also extended by an additional drowning child

scenario, the psychophysical numbing scenario. My findings offer first suggestions that this scenario might be worth including in the Drowning Child Argument when it is used in research and practice.

The results show that this moral argument has a strong effect on whether people report it is wrong not to donate. This was widely believed anecdotally, as the argument has been used by different people in numerous occasions and confirmed in a first study by Caviola (manuscript in preparation) but it is valuable to have some more evidence from a randomised control study about this. It is too early to draw strong conclusions, but my findings suggest that defensiveness plays a role in how well the Drowning Child Argument is received. If defensiveness plays a role in accepting the widely used Drowning Child Argument for donating, then it could make sense to self-affirm people before presenting them with the argument. This could be interesting for anyone using this or similar thought experiments, whether on websites, in workshops, classes, talks, or maybe even in private conversation. The mechanisms through which self-affirmation works to decrease defensiveness are a worthwhile field of further study and would benefit from an overarching, unifying framework. The more we know about how to decrease defensiveness, the better we can use this knowledge to set the context for people so that they can approach negative but important information openly. This can enable us to make better decisions that are more in tune with our own values.

References

- Andrade, E. B., & Ariely, D. (2009). The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes*, 109, 1–8.
- Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion*, 22, 1–32.
- Blanton, H., Cooper, J., Skurnik, I., & Aronson, J. (1997). When bad things happen to good feedback: Exacerbating the need for self-justification with self-affirmations. *Personality and Social Psychology Bulletin*, 23, 684–692.
- Brown, J. D. (1990). Evaluating one's abilities: Shortcuts and stumbling blocks on the road to self-knowledge. *Journal of Personality and Social Psychology*, 50, 149–167.
- Caviola, L. *Rationalizations in Morally Demanding Cases: Perceptions of Singer's Drowning Child Argument*. Manuscript in preparation.
- Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin*, 26, 1151–1164.
- Comer, D. R., & Sekerka, L. E. (2017). Keep calm and carry on (ethically): Durable moral courage in the workplace. *Human Resource Management Review*, In Press, Corrected Proof. doi:10.1016/j.hrmr.2017.05.011
- Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology*, 40, 350–356.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96, 608–630.

- Crocker, J., Niiya, Y., & Mischkowski, D. (2008). Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive other-directed feelings. *Psychological Science*, 19(7), 740–7. doi:10.1111/j.1467-9280.2008.02150.x
- Dunning, D. (1995). Trait importance and modifiability as factors influencing self-assessment and self-enhancement motives. *Personality and Social Psychology Bulletin*, 21, 1297–1306.
- Ehrlich, G. A., & Gramzow, R. H. (2015). The politics of affirmation theory: When group-affirmation leads to greater ingroup bias. *Personality & Social Psychology Bulletin*, 41(8), 1110–22. doi:10.1177/0146167215590986
- Epton, T., Harris, P. R., Kane, R., van Koningsbruggen, R., & Sheeran, P. (2015). The impact of self-affirmation on health-behavior change: A meta-analysis. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, 34(3), 187–96. doi:10.1037/hea0000116
- Fein, S., & Spencer, S. J. (1997). Prejudice as Self-Image Maintenance: Affirming the Self Through Derogating Others. *Journal of Personality and Social Psychology*, 73(1), 31–44.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. California: Stanford University Press.
- Fetherstonhaugh, D., Slovic, P., Johnson, S., & Friedrich, J. (1997). Insensitivity to the value of human life: A study of psychophysical numbing. *Journal of Risk and Uncertainty*, 14(3), 283–300. doi:10.1023/A:1007744326393
- Forgas, J., & Tan, H. (2013). To give or to keep? affective influences on selfishness and fairness in computer-mediated interactions in the dictator game and the ultimatum game. *Computers in Human Behavior*, 29(1), 64–74. doi:10.1016/j.chb.2012.07.017

- Gaudine, A., & Thorne, L. (2001). Emotions and ethical decision-making in organizations. *Journal of Business Ethics*, 31(2), 175–187.
- Gerrards-Hesse, A., Spies, K., & Hesse, F. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology*, 85(1), 55–78. doi:10.1111/j.2044-8295.1994.tb02508.x
- Halmburger, A., Baumert, A., & Schmitt, M. (2015). Anger as driving factor of moral courage in comparison with guilt and global mood: A multimethod approach. *European Journal of Social Psychology*, 45(1), 39–51.
- Harlé, K. M., & Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the ultimatum game. *Emotion (Washington, D.C.)*, 7(4), 876–81.
- Harvey, R., & Oswald, D. (2000). Collective guilt and shame as motivation for white support of black programs. *Journal of Applied Social Psychology*, 30(9), 1790–1811.
- Isen, A. M., & Levine, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology*, 21, 384–388.
- Ipeirotis, P. (2009). *Turker Demographics vs Internet Demographics*. Retrieved from <http://www.behind-the-enemy-lines.com/2009/03/turker-demographics-vs-internet.html> on 19th August 2017.
- Ipeirotis, P. (2010). *The New Demographics of Mechanical Turk*. Retrieved from <http://hdl.handle.net/2451/29585> on 19th August 2017.
- Jacks, J. Z., & O'Brien, M. E. (2004). Decreasing resistance by affirming the self. In E. S. Knowles and J. A. Linn (Eds.), *Resistance and persuasion* (pp. 235–257). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kohlberg, Lawrence (1971). *From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development*. New York: Academic Press.

- Koole, S. L., Smeets, K., Van Knippenberg, A., & Dijksterhuis, A. (1999). The Cessation of Rumination Through Self-Affirmation. *Journal of Personality and Social Psychology*, 77(1), 111–124.
- McQueen, A., & Klein, W. M. P. (2006). Experimental manipulations of self-affirmation: A systematic review. *Self and Identity*, 5, 289–354.
- Niesta Kayser, D., Greitemeyer, T., Fischer, P., & Frey, D. (2010). Why mood affects help giving, but not moral courage: Comparing two types of prosocial behaviour. *European Journal of Social Psychology*, 40(7), 1136–1157. doi:10.1002/ejsp.717
- North, A. C., Tarrant, M., & Hargreaves, D. I. (2004). The effects of music on helping behaviour. A field study. *Environment and Behavior*, 36, 266–275.
- Olejnuk, A. B., & LaRue, A. A. (1980). Affect and moral reasoning. *Social Behavior and Personality*, 8, 75–79.
- Pagano, V., & Debono, K. (2011). The effects of mood on moral judgment: The role of self-monitoring. *Journal of Applied Social Psychology*, 41(12), 2928–2942. doi:10.1111/j.1559-1816.2011.00864.x
- Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, 56, 365–392.
- Piper, G., & Schnepf, S. (2008). Gender differences in charitable giving in Great Britain. *Voluntas: International Journal of Voluntary and Nonprofit Organizations : Official Journal of the International Society for Third-Sector Research*, 19(2), 103–124. doi:10.1007/s11266-008-9057-9
- Quinn, N. (2006). The self. *Anthropological Theory*, 6(3), 362–384.

- Raghunathan, R., & Trope, Y. (2001). Walking the Tightrope between Feeling Good and Being Accurate: Mood-as-a-Resource in Processing Persuasive Messages. *Advances in consumer research*, 28, 114.
- Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, 67, 525–539.
- Rest, J. R. (1975). Longitudinal study of the Defining Issues Test of moral judgment: A strategy for analyzing developmental change. *Developmental Psychology*, 11, 738–748.
- Schmeichel, B. J., & Martens, A. (2016). Self-Affirmation and Mortality Salience: Affirming Values Reduces Worldview Defense and Death-Thought Accessibility. *Personality and Social Psychology Bulletin*, 31(5), 658–667.
- Sherman, D., & Cohen, G. (2006). The Psychology of Self-defense: Self-Affirmation Theory. *Advances in Experimental Social Psychology*, 38, 183–242.
- Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do Messages about Health Risks Threaten the Self? Increasing the Acceptance of Threatening Health Messages Via Self-Affirmation. *Personality and Social Psychology Bulletin*, 26(9), 1046–1058.
- Singer, P. (1972). Famine, Affluence, and Morality. *Philosophy and Public Affairs*, 1(3), 229–243. doi:10.2307/2265052
- Singer, P. (2010). The life you can save: *Acting now to end world poverty*. Melbourne: Text Publishing Company.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 21, pp. 261–302). New York: Academic Press.

- Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, 45(1), 5–19.
- Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: the role of affirmational resources. *Journal of Personality and Social Psychology*, 64(6), 885–896.
- Stone, J., Whitehead, J., Schmader, T., & Focella, E. (2011). Thanks for asking: Self-affirming questions reduce backlash when stigmatized targets confront prejudice. *Journal of Experimental Social Psychology*, 47(3), 589–598.
doi:10.1016/j.jesp.2010.12.016
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171.
- Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, 4, 290–299.
- Thompson, E. R. (2016). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, 38(2), 227–242.
- Trope, Y. (1975). Seeking information about one's own ability as a determinant of choice among tasks. *Journal of Personality and Social Psychology*, 32, 1004–1013.
- Trope, Y., & Fishbach, A. (2000). Counteractive self-control in overcoming temptation. *Journal of Personality and Social Psychology*, 79, 493–506.
- Trope, Y., & Neter, E. (1994). Reconciling competing motives in self-evaluation: The role of self-control in feedback seeking. *Journal of Personality and Social Psychology*, 66, 646–657.
- United Nations Inter-agency Group for Child Mortality Estimation (2015). Levels and trends in child mortality: Report 2015. Retrieved from

http://www.childmortality.org/files_v20/download/IGME%20Report%202015_9_3%20LR%20Web.pdf

- Wang, X., & Zhao, X. (2017). The mediating role of temporal considerations on the effects of self-affirmation on responses to organ donation messages. *Health Communication, 1*–8, 1–8. doi:10.1080/10410236.2016.1250190
- Wellen, J. M., Hogg, M. A., & Terry, D. J. (1998). Group norms and attitude-behaviour consistency: The role of group salience and mood. *Group Dynamics, 2*, 48–56.
- Zarinpoush, F., Cooper, M., & Moylan, S. (2000). The effects of happiness and sadness on moral reasoning. *Journal of Moral Education, 29*, 397–412.

Appendix

Appendix A: Drowning Child Argument Condition

Now some questions will follow related to a thought experiment called the drowning child scenario.

There are many situations in which we are in a position to help someone in need, and have to decide whether to help them or to do nothing. We will now describe one such situation.

You see a small child drowning in a shallow pond. You could walk into the pond and save the child's life, but doing so would ruin your \$500 suit. You are faced with the decision of whether you want to save the child's life at the expense of \$500 or keep the money to yourself and let the child die.

'If we can easily save the life of an innocent person who is in danger at little cost to ourselves, it would be morally wrong for us to do nothing.'

This seems right. After all, what is \$500 compared to the life of a child?

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

'Even if there are other people who could help the person, but they do nothing, it would still be morally wrong for us to just watch and not help if we can.'

This seems right. After all, if other people aren't doing what they should, why should this excuse us from doing what we ought to do? Someone has to save the drowning child. If other people aren't helping, then surely we must do so.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

'Even if there is a small chance that we won't succeed in saving the life of the person in danger, it would still be morally wrong for us to do nothing.'

This seems right. After all, someone's life is at stake, and as long as there is a good enough chance that we will succeed in saving them, it's surely wrong not to even try.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

'Even if the person in danger is physically far away from us, it would still be morally wrong for us to do nothing if we could easily help.'

This seems right. After all, if you can save a life, why should it matter how far away someone is from us? If, for example, we found out that a child in another city is about to drown in a pond, and we could easily make a phone call to tell this to someone in that city who could then rescue the child, surely it would be wrong for us not to make the phone call.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

'Even if there were a hundred children drowning but we were only able to save one, it would still be morally wrong for us to do nothing.'

This seems right. After all, this child's life is worth saving independently of other lives that are worth saving too.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

Do you think that if you and your friends were born in a very poor country, receiving money from someone much better off could make a difference in your community? What might you use the money for?

Now, the following argument can be made:

1. Millions of people around the world live in extreme poverty and are dying due to hunger and disease. These people need our help just like the child that is drowning in the pond.
2. There are effective charity organizations that have been scientifically proven to save the lives of people in need using even small amounts of money received through donation such as \$500. E.g.
 - The organisation GiveDirectly directly transfers donations to the world's poorest people. “Strong evidence indicates that cash transfers lead recipients to spend more on their basic needs (such as food) and may allow recipients to make investments with very high returns, with no evidence of large increases in spending on items like alcohol or tobacco (for more, see our full report on cash transfers). (1)
 - Malaria is a disease transmitted through mosquito bites. The Against Malaria Foundation distributes insecticide-treated malaria bednets for 5 dollars per net. Areas in which this intervention was implemented saw a steep decline in malaria infections. (1) “Malaria is not just a disease commonly associated with poverty: some evidence suggests that it is also a cause of poverty and a major hindrance to economic development. (2) (3) (...) Cerebral malaria is one of the leading causes of neurological disabilities in African children.” (4)

- Many more organisations have been scientifically proven to save and improve people's lives and lift communities out of disease and poverty, cheaply in developing countries and also provenly but less cheaply in developed countries.
3. Just like in the situation when we can easily save the drowning child, we are in a position to help at relatively little cost to ourselves. By donating money to such charities, we are helping people who will otherwise die or suffer unnecessarily.

Conclusion:

Therefore it would be morally wrong for us NOT to donate some of our money to effective charities that can save lives. In fact, one could argue that it would be similarly wrong for us not to donate as it would be to not save the drowning child.

References

- (1) <http://www.givewell.org/effective-altruism>. GiveWell is a charity evaluator.
- (2) Gollin D, Zimmermann C (2007). Malaria: Disease Impacts and Long-Run Income Differences (PDF) (Report). *Institute for the Study of Labor*.
- (3) Worrall E, Basu S, Hanson K (2005). "Is malaria a disease of poverty? A review of the literature". *Tropical Health and Medicine*. 10 (10), 1047–59.
- (4) Idro R, Marsh K, John CC, Newton CR (2010). "Cerebral malaria: Mechanisms of brain injury and strategies for improved neuro-cognitive outcome". *Pediatric Research*. 68 (4), 267–74. doi:10.1203/PDR.0b013e3181eee738. PMC 3056312 Freely accessible. PMID 20606600

Appendix B: Drowning Child Argument Control Condition

Please read the following text and answer the questions:

Imagine you come across a small child who is playing in the sand. You could play along, but you have just come from a meeting and are wearing an expensive pair of shoes and a \$500

suit that will be ruined forever if you do. This child lives in your neighbourhood, so you will have other opportunities to play with her. Other adults standing around also do not join the child, instead they relax and chat with each other. The child looks happy playing on her own. 'If this was a well-behaved child I know who would love to play with me and I was wearing clothes that I did not mind getting dirty, it would be enjoyable to play along.

This seems right. After all, who doesn't like spending a bit of time with well-behaved children who are happy to see you?

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

'Most small children enjoy playing in sand.'

This seems right. After all, small children tend to love to play anywhere and like to be creative, which they can be with sand.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

‘The child is likely to be even happier if she had friendly company playing with her.

This seems right. After all, she has been playing by herself for a while already and company is fun and she could show what she has built so far.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

‘Children should learn to play happily on their own.’

This seems right. After all, it is good for children to learn to be independent.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

It is not worth ruining my shoes and suit to play with this child.

This seems right. After all, there are many better moments than this one to play along and the child seems happy enough.

Strongly agree

Agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Disagree

Strongly disagree

What might the child be building out of the sand and what toys might she be using?

In fact, while it might be more fun for the child to have someone interacting with her, there are potential benefits of playing by herself. The following argument can be made (extract from thespruce.com by Apryl Duncan):

Playing alone:

1. Teaches children to have fun independently (...) As your children grow, they understand that they won't always have someone by their side every waking moment. They'll be more confident and satisfied individuals.

2. Brings out their imagination

You may already feel like your child is full of imagination. Just wait until you step back and let them play by themselves. The time alone will draw out superheroes, princesses and other play situations you wouldn't get to see if they weren't engaging in solo playtime. They'll be quick to think on their feet and their creativity will really shine. (...)

3. Encourages calmness

Playing by themselves brings a sense of calmness to your kids. This time playing alone takes their mood to a different level as they peacefully play with their toys. (...)

4. Gets kids ready for school

For preschoolers, you're probably the number-one playmate your children have ever known. As you back away and show them how to play by themselves, they understand that you're not always physically going to be there with them. (...)

References

<https://www.thespruce.com/why-playing-alone-is-important-3129415>

Damast, A. (1996). Stability and prediction in children's solitary play and attention across the first two years. *Infant Behavior and Development: Supplement 1*, 19, 415-415.

doi:10.1016/S0163-6383(96)90469-5

Lloyd, B. (2003). Solitary play and convergent and divergent thinking skills in preschool children. *Early Childhood Research Quarterly*, 18(1), 22-41.

doi:10.1016/S0885-2006(03)00004-8

Appendix C: Rosenberg self-esteem scale

1. On the whole, I am satisfied with myself.
2. At times I think I am no good at all. (R)
3. I feel that I have a number of good qualities.
4. I am able to do things as well as most other people.
5. I feel I do not have much to be proud of. (R)
6. I certainly feel useless at times. (R)
7. I feel that I'm a person of worth, at least on an equal plane with others.

8. I wish I could have more respect for myself. (R)
9. All in all, I am inclined to feel that I am a failure. (R)
10. I take a positive attitude toward myself.

Items 2, 5, 6, 8, and 9 were reverse coded.