

Running head: Dark humor and moral judgment

Title: When dark humor and moral judgment meet in sacrificial dilemmas: Preliminary evidence with females

Emmanuelle BRIGAUD & Nathalie BLANC

University Paul Valéry Montpellier 3

Epsilon EA 4556

Abstract

The influence of dark humor on moral judgment has never been explored, even though this form of humor is well-known to push the boundaries of social norms. In the present study, we examined whether the presence of dark humor leads female participants to approve a utilitarian response (i.e., to kill one to save many) in sacrificial dilemmas. The effects of two types of humorous contexts were compared (i.e., dark versus nondark) on dilemmas, which differed according to whom benefits from the crime (i.e., oneself and others versus others only). In addition to collecting moral responses, individuals' emotional states were assessed at three critical steps: Before and after reading the jokes and also after performing the moral judgment task. Our results revealed that dark and nondark humor similarly elicited a positive emotional state. However, dark humor increased the permissiveness of the moral violation when this violation created benefits for oneself. In self and other beneficial dilemmas, female participants in the dark humorous condition judged the utilitarian response more appropriate than those in the nondark condition. This study represents a first attempt in deepening our understanding of the context-dependent nature of moral judgment usually assessed in sacrificial dilemmas.

Keywords: moral judgment, sacrificial dilemmas, dark and nondark humor

1. Introduction

Suppose a runaway trolley is about to run over and kill five people. Suppose further that a large stranger is standing on a bridge over the tracks and that the only way to stop the trolley is to push that person in front of the trolley, killing him for sure but saving the others. Would it be okay to sacrifice one life to save several others? Most people answer “no” to this “high conflict” personal moral dilemma (Greene et al., 2001, 2004).

This phenomenon has been widely studied by psychologists to understand the cognitive and affective processes underlying moral judgments (see Waldman, Nagel, & Wiegmann, 2012 for a review; see also Bartels, Bauman, Cushman, Pizarro, & McGraw, 2015). The dual-process theory provides a relevant framework to explain people’s responses to sacrificial dilemmas (Greene, 2007; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). According to this well-known theory, two separate systems are involved in moral judgment: the controlled cognitive process, which corresponds to conscious reasoning (slow and effortful), and the automatic emotional one based on intuition and affects (fast and largely unconscious). In response to “high conflict” personal dilemmas, like the footbridge scenario described above (Thomson, 1985), people are typically driven by automatic emotional responses and judge that it is morally unacceptable to push someone off a footbridge even though not pushing him would result in a greater number of deaths. The perspective entailing a moral violation, such as killing an innocent person triggers a strong emotional aversion that inhibits an amoral solution (Greene, 2008). However, with sufficient time, motivation or resource conditions, people may engage in controlled cognitive processes regarding the costs and the benefits of killing another person. Such mechanisms result in a utilitarian judgment: approving the sacrifice of one life in order to save the lives of five is morally acceptable with respect to the number of victims.

In a footbridge-type dilemma, the rational choice (i.e., sacrifice the life of one person in order to save the lives of a greater number of people) is directly in conflict with deontological rules or intuition. Consequently, the utilitarian response requires one to control or overcome the initial aversive reaction against harming an innocent person (Greene, 2008). Two sets of studies have provided converging evidence in line with this idea. First, both empirical and neuropsychological studies have shown that utilitarian judgments are due to an absent or a reduced affective response. Participants with emotion-related neurological deficits (e.g., Ciaramelli et al., 2007; Koenigs et al., 2007; Moretto et al., 2009), with decreased empathic concern or with antisocial personality traits (Bartels & Pizarro, 2011; Conway & Gawronski, 2013; Gleichgerrcht & Young, 2013; Kahane et al., 2015) reach more utilitarian conclusions in sacrificial dilemmas. Secondly, other researchers have shown a link between the utilitarian response and working memory capacity (Moore, Clark, & Kane, 2008) and also between this type of response and an individuals' need for cognition (Bartels, 2008; Conway, & Gawronski, 2013). Precisely, participants are more likely to approve a utilitarian response when they scored high in working memory capacity or in need for cognition (a motivational tendency to seek and enjoy effortful cognitive activity).

Concomitantly, environment-induced positive mood at the time of judgment increases a utilitarian response. For instance, simple exposure to humorous material before the presentation of the footbridge scenario increases permissiveness for moral violations (i.e., pushing the stranger over the bridge) (Strohmingher, Lewis, & Meyer, 2011; Valdesolo & DeSteno, 2006). Such effect arises, because humor is usually associated with the experience of a positive emotion (i.e., mirth, laughter, pleasure). Therefore, if humor induces positive emotion at the time of judgment, the perceived negativity linked to any potential moral violation is attenuated and, thus, utilitarian response increases. This explanation is consistent

with Fredrickson's (2001) hypothesis that positive emotions can act as an antidote to negative emotions correcting or diminishing their influence.

Humor, used as a communicative activity, elicits positive emotional reactions in perceivers and tendency to laugh (Gervais & Wilson, 2005; Martin & Ford, 2018; Veatch, 1998). It also indicates to the target or audience that what happening, or is going to happen, should be taken as a joke (Gervais & Wilson, 2005; Ramachandran, 1998). In Ford, Boxer, Armstrong and Edel's (2008) words: "humor invokes a conversational rule of levity, that is, humor communicates an implicit message to the receiver that the usual rules of logic and expectations of common sense did not apply" (p. 160). In the context of social judgment, this central property of humor might allow us to treat the violation of moral rules (e.g., "it's forbidden to kill") as a matter of play and, therefore, favor the utilitarian judgment. Consistent with this hypothesis, Strohminger, Lewis and Meyer (2011) found that mirth (i.e., the positive emotion associated with humor) increases permissiveness for deontological violation in moral dilemmas, whereas elevation (i.e., a positive emotion experienced upon witnessing another person perform a virtuous act; Algie & Haidt, 2009; Haidt, 2003) has the opposite effect. This result highlighted that the influence of humor on people's moral judgment cannot be explained simply in terms of experiencing positive emotions. They suggest that humor influences moral judgment by removing the gravitas of the moral violation (i.e., making immoral behavior funny). This interpretation is consistent with the Benign Violation Theory of humor (McGraw & Warren, 2010; McGraw, Warren, Williams, & Leonard, 2012; Veatch, 1998), which suggests that humor occurs when people simultaneously appraise a violation as being normal, acceptable, or okay.

The aim of the present study was to investigate more thoroughly to what extent a humorous context can influence the response of participants in personal moral dilemmas. The

only two studies (Valdesolo & DeSteno, 2006; Strohming, Lewis, & Meyer, 2011) that focused on the role of humor on moral judgment used non-transgressive humor (i.e., inoffensive comedy). Thus, it would be interesting to see if the observed humor effect on moral judgment could be stronger when one uses humor with a transgressive content. As this form of humor is closely linked to the transgression of social norms and moral systems, it creates a context that could lead individuals to consider the utilitarian response in sacrificial dilemmas as harmless or okay. McGraw and Warren (2010) showed that moral violation is benign when another norm suggests that the behavior is acceptable or correct. In this sense, expression of transgression delivered in a humorous form could suggest that, in this context, the moral violation is okay (i.e., it's acceptable to kill someone). This is closely linked with recent research that has shown that exposure to humorous forms of disparagement (i.e., sexist, racist or anti-gay jokes) lead to an increase in expression of prejudice toward target groups (e.g., O'Connor, Ford, & Banos, 2017; Saucier, O'Dea, & Strain, 2016; Woodzicka & Ford, 2010). According to the benign-violation theory hypothesis, such effects may occur because in the humorous context, the moral violation (i.e., denigrate a social group) is perceived as benign (see Gutiérrez, Carretero-Dios, Willis, & Morales, 2018; Thai, Borgella, & Sanchez, 2019, for a similar interpretation).

A particularly interesting form of humor to study in sacrificial dilemmas is dark humor because it treats sinister and tragic subjects, like death, with amusement and trivializes the victim's suffering (Aillaud & Piolat, 2012). Dark humor (death-related humor) is described as cynical, gallows, morbid. Playing with serious or sad real life events, dark humor is generally considered as transgressive since it crosses the red line of social norms and moral systems. This form of humor takes its name from jokes about condemned men or hopeless victims to relieve tension before being executed (Freud, 1905). Thus, the condemned to death may well declare when led to the scaffold on a beautiful Monday morning, "This is a week

that is starting well!" The juxtaposition of morbid and farcical elements in dark humor elicits two simultaneous perceptions: one, that the dark content constitutes a moral violation in which negative serious topics are mocked and, the other, an interpretation that the dark content is benign. Since dark humor treats negative serious ideas (like death, suffering of the victims or body integrity) in a light-hearted, playful manner (Chareaudeau, 2006), it is not surprising that people use it as a coping strategy in stressful life-and-death situations (Christopher, 2015; Rowe & Regehr, 2010; van Wormer & Boes, 1997; Young, 1995). In the context of social judgment, dark humor might signal that the violation of moral rule (i.e., to kill someone) is fine and, therefore favor the utilitarian response. Thus, we predicted that participants exposed to dark humor before performing a moral judgment task would answer in a more utilitarian fashion (i.e., approving killing one) compared to participants exposed to nondark humor.

To deepen our understanding of moral judgment in a humorous context, we questioned whether the humor effect depends on who benefits from the crime. Our research considered cases where crime benefits oneself and others versus others only. We predicted that dark humor reinforces the tendency to morally approve the act of killing someone in particular when such action benefits the self in addition to others. This prediction is consistent with two results in psychology of morality (for a review see Ellemers, van der Toorn, Paunov, & van Leeuwen, 2019). The first one showed that the tendency to produce utilitarian responses is strongly tied to consideration of self-interest (Christensen, Flexas, Calabrese, Gut, & Gomila, 2014; Kahane, Everett, Earp, Farias, & Savulescu, 2015). Indeed, people are more inclined to approve harm if their own life is at stake than if the moral transgression is merely to save others. The second one suggested that people feel less negative emotions (e.g., guilt and shame) about their dishonest actions and perceived them to be morally acceptable when there are other beneficiaries for these actions in addition to themselves (Gino, Ayala, &

Ariely, 2013). In this case, people use the potential benefits for others as a way to justify their self-interested and unethical actions (self-serving altruism). After being exposed to dark humor, committing harm could be considered less socially inappropriate when participants are faced with self and other beneficial dilemmas, because the crime relies on their own utility since this “selfish” consideration enables them to save others too.

The objective of the present research was to investigate more specifically the impact of two humorous contexts (i.e., dark and nondark jokes) on people’s responses to sacrificial dilemmas) as a function of whom benefits from the crime (other vs. self and other).

2. Method

2.1 Participants

One hundred and thirty-six female undergraduate students¹ from the University of Montpellier 3 (France) took part in this experiment. The average age of the sample was 20.75 years ($SD = 3.40$). Informed consent was obtained from all students prior to participating in any of the tasks. They were informed that their responses remained anonymous in respect of the Data Protection law. All students received course credit as compensation.

2.2 Materials

2.2.1 Humorous materials. To assess the effects of humor on moral judgment, we used 12 jokes: six jokes were not transgressive (i.e., nonsense or clownish humor) and six dark jokes with a transgressive content (i.e., dealing with sinister topics with amusement like death, suffering of the victims and body integrity). For example, one of the dark jokes used:

¹- The original sample was composed of 136 female and 19 male participants (which is a classical sample of students enrolled in psychology courses). Because there are gender differences in the perception and the attitude toward dark and nondark humor (e.g., Aillaud & Piolat, 2012, 2014), we selected female participants only for this study.

A lawyer goes to the coroner about an autopsy:

- Before signing the death certificate, did you take this man’s pulse?
- No.
- Did you check to see if his heart was still beating?
- No.
- Did you check whether he was still breathing?
- No.
- So you signed this death certificate without performing any of the recommended tests for establishing whether a person is really and truly dead?
- Yes. Why? Did you find his head?

These jokes were selected on the basis of pretest ratings given by a total of 180 undergraduate students from the University of Montpellier 3, France (M age = 19.84 years, SD = 2.63; the majority were female, 86%). They were all volunteers and were compensated with course credit for their participation.

A first group of 90 participants were asked to rate 30 jokes regarding their darkness. They rated “How dark is the joke?” using a scale from 1 (not at all dark) to 5 (very dark). Because participants were tested collectively, jokes were presented in a counterbalanced order across participants. Based on the results of this pilot study, we selected 12 jokes from the pool of the 30 rated jokes (i.e., those that lead to the most consistent appraisal among the sample): 6 jokes were attributed the lowest score (i.e., 1 = “*not at all dark*”) by at least 78% of the sample and 6 jokes were attributed the highest score (i.e., 5 = “*very dark*”) by at least 72% of the sample. All other jokes were excluded from the experimental material.

To ensure that this set of jokes (i.e., dark and nondark jokes) was similar in terms of funniness ratings, but distinct in terms of transgressive content, we recruited a second group of 90

participants. After reading each joke, they answered the two following questions: “How funny is the situation described in this joke?” and “How unbecoming and unseemly² is the situation described in this joke?”. Using the same procedure as Aillaud and Piolat (2012, 2014), responses were made on a 4-point scale (1- “*definitely not*”, 2- “*not*”, 3- “*slightly yes*”, and 4- “*definitely yes*”). Note that this 4-point scale enabled us to avoid a midpoint evaluation.

A two-way ANOVA (Type of humor: Dark vs. Nondark) was made for each rating. These analyses revealed a main effect of transgressive content ratings only, $F(1, 88) = 77.45, p < .001, \eta_p^2 = .47$, dark jokes being judged as more unbecoming and unseemly ($M = 2.64, SD = .69$) than nondark ones ($M = 1.23, SD = .32$). There were no significant differences between dark and nondark jokes regarding funniness ($M = 2.57, SD = .48$ and $M = 2.69, SD = .63$, respectively), $F(1, 88) = .93, p = .34, ns$. These results confirm that participants perceived a difference between dark and nondark humor solely on the transgressive dimension.

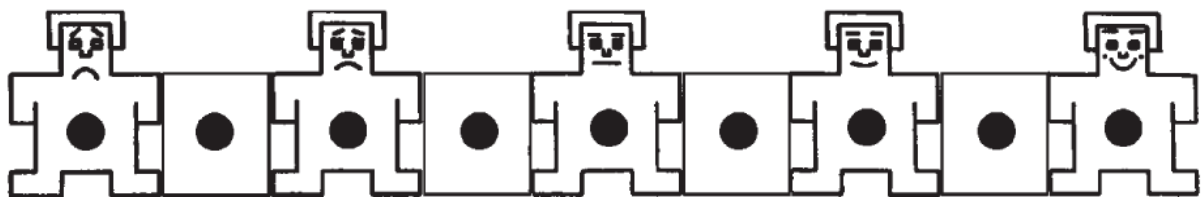
2.2.2 Moral dilemmas. We selected four high conflict personal dilemmas from a previously used set (see Greene et al., 2001, 2004) in which the participant was always presented as the main protagonist of the situation (i.e., the one who was supposed to carry out the moral violation). The dilemmas were similar regarding at least two dimensions: All dilemmas involved killing one person in order to save several others; the number of people saved was comparable ($N = 10$). In addition, all these dilemmas were known to elicit mainly the same negative emotion (i.e., guilt) during judgment (Choe & Min, 2011). The dilemmas were only distinguished according to whom benefits (other vs. self and other) from the crime. The footbridge and the vitamins were the two other-beneficial dilemmas, while the lifeboat and safari were self and other beneficial ones. In the latter, the crime enabled one to save others as well as the protagonist herself. For example, in the lifeboat dilemma, the protagonist

² Indecent, inappropriate

must choose whether to throw a person overboard to save the life of remaining passengers and her life too.

2.2.3 Emotional scales. Participant's emotional state was assessed on two dimensions: valence (positive vs. negative) and arousal (level of activation) using the Valence and the Arousal scales of the Self-Assessment Manikin (SAM, Lang, 1980). According to Bynion and Feldner (2017), the SAM is a brief and nonverbal measure of emotional state which reliability has been confirmed by numerous studies conducted in various domains (e.g., psychology, communication, advertising; Morris, 1995) and populations (e.g., gender, age, race; Backs, da Silva, & Han, 2005; Nabizadeh Chianeh, Vahedi, Rostami, & Nazari, 2012). The SAM scales consist of two sets of five figures depicting different levels of affective valence and arousal (see Figure 1). For each dimension, participants were instructed to place an "X" on or between the figures that best described their emotional state. The Valence scale (a) ranged from unhappy (1) to happy (9) and the Arousal scale (b) from calm (1) to excited (9).

(a)



(b)

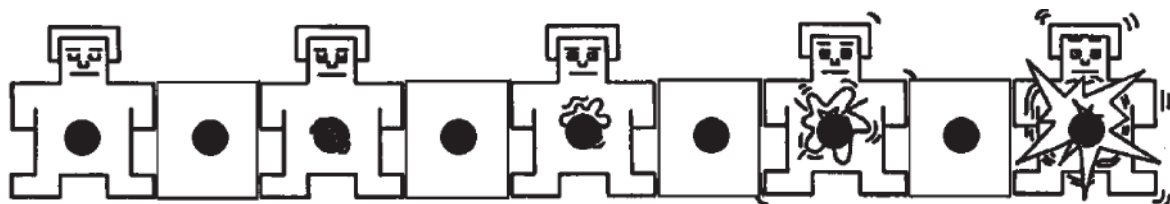


Figure 1. The SAM Scales for valence (a) and arousal (b) with the five figures and the spaces between them corresponding to one point of the nine-point scale (Lang, 1980; see also Lang, Bradley, & Cuthbert, 1997)

2.3 Procedure

After giving their informed consent, participants were randomly assigned either to the dark humor condition ($N = 68$) or to the nondark humor condition ($N = 68$). All participants were asked to complete an online questionnaire composed of two parts: First, they were exposed to six jokes and then they had to complete a moral judgment task. Immediately after reading the humorous material, all participants responded to four high-conflict personal dilemmas. The order of presentation of the dilemmas was counterbalanced within and between the dark and nondark humor conditions. For each dilemma, participants had to decide whether the utilitarian option (i.e., to kill someone) was appropriate or not (yes/no question). The answer “yes” always represented the utilitarian response. The scenarios were briefly introduced by stating that they refer to serious situations that could be seen as unpleasant but require making a difficult choice. To put the participants “in context” for the task that awaited them, and to ensure that they were engaged in the moral issues at stake, they were instructed to imagine themselves in each situation so that their answer could mirror their action in real life (e.g., see Tassy, Oullier, Mancy, & Wicker, 2013). They were asked to be as honest as possible in their responses, knowing that there is no good or wrong answer. In addition to recording responses to the dilemmas, we also assessed the participant’s emotional state in three steps: before reading the jokes (Time 1), after reading the jokes (Time 2) and after the moral judgment task (Time 3). Participants rated their emotional state using the Valence and Arousal scales of the Self-Assessment Manikin (SAM, Lang, 1980).

3. Results

3.1 Emotional states. To examine whether participants' emotional states fluctuated throughout the experiment, a repeated analysis of variance (ANOVA) was conducted, first on the valence ratings and, second on arousal ratings. The type of humor (i.e., Nondark humor, Dark humor) was the between-participant factor, and the mood assessment time (Time 1, Time 2, Time 3) was the within-participant factor. Mean ratings (and standard deviation) of valence and arousal are reported in Table 1.

Table 1. *Mean ratings (and standard deviations) of valence and arousal for the dark humor and the nondark humor conditions at each of the three assessment times.*

	Valence ratings			Arousal ratings		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
Nondark humor	5.37 (1.88)	6.13 (1.65)	4.06 (2.04)	4.75 (1.82)	5.19 (1.85)	5.51 (2.09)
Dark humor	5.50 (1.76)	5.97 (1.71)	4.28 (1.80)	5.06 (1.91)	5.19 (1.97)	5.47 (2.06)

Notes. Before (Time 1) and after (Time 2) reading the jokes; after the moral judgment task (Time 3)

Regarding valence ratings, a significant effect of time assessment was observed, $F(2, 268) = 59.34$, $p < .001$, $\eta_p^2 = .31$. Post hoc analysis (Scheffé test) revealed that participants reported feeling happier after reading the jokes than before reading the jokes (Time 2, $M = 6.05$, $SD = 1.68$; Time 1, $M = 5.43$, $SD = 1.82$), but their induced happiness decreased after the moral judgment task (Time 3, $M = 4.17$, $SD = 1.92$) ($ps < .001$). Regarding arousal ratings, ANOVA revealed a significant effect of time assessment, $F(2, 268) = 5.51$, $p < .01$, $\eta_p^2 = .04$. Participants reported feeling more excited after (Time 3, $M = 5.49$, $SD = 2.07$) than before the experiment (Time 1, $M = 4.90$, $SD = 1.86$), ($p < .01$). No other effects were significant.

3.2 Moral judgment. The mean proportion of utilitarian responses (i.e., killing is judged morally appropriate) was analyzed (ANOVA) to explore the effect of both the type of humor (i.e., Nondark humor, Dark humor) and the type of dilemma (i.e., Self and Other-beneficial vs. Other-beneficial). This 2 x 2 analysis showed a main effect of the type of dilemma indicating that killing to save oneself and others was judged to be more appropriate ($M = .55$, $SD = .38$) than killing to save only others ($M = .19$, $SD = .30$), $F(1, 134) = 94.99$, $p < .001$, $\eta_p^2 = .41$. A significant Type of humor x Type of dilemma interaction showed that this tendency to accept moral violation in their own self-interest increased when participants were exposed to dark jokes, $F(1, 134) = 7.75$, $p < .006$, $\eta_p^2 = .05$ (see Figure 2). Post hoc analysis (Scheffé test) revealed that, in self and other beneficial dilemmas, the mean proportion of utilitarian responses was significantly higher in the dark humor condition than in the nondark ones ($p < .001$). No significant difference between these two conditions was found for other-beneficial dilemmas ($p = .60$). No other effects were significant.

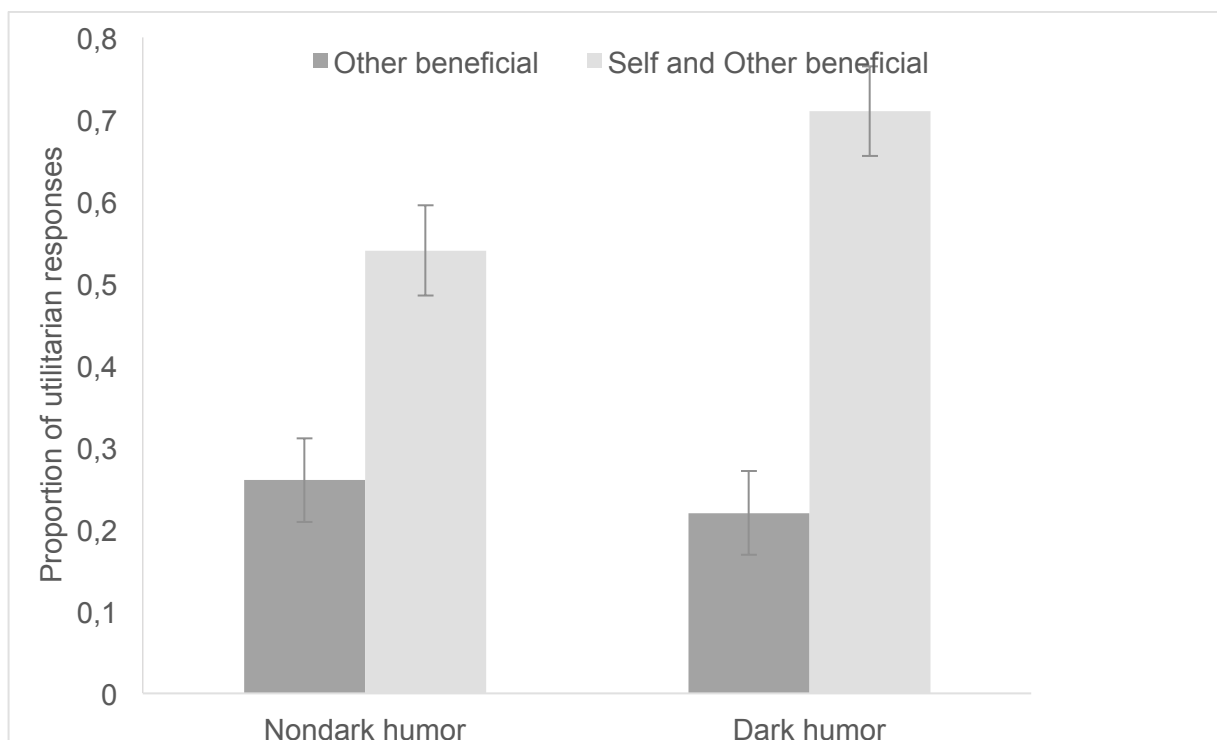


Figure 2. Mean proportion of utilitarian responses as a function of Type of humor (dark vs. nondark) and Type of dilemma (self and other beneficial vs. other beneficial). Error bars depict standard errors.

4. Discussion

Compared to the existing literature dealing with humor and moral judgment (Valdesolo & DeSteno, 2006; Strohminger et al., 2011), the present study highlights the relevance of distinguishing different types of humorous inducing materials. Contrary to previous studies that only considered nondark humor, our comparison between dark and nondark humor revealed that variability in moral judgment could not simply be explained in terms of experiencing positive emotions or in terms of the levity property of humor. The content of humorous jokes deserves to be considered especially when this content relies on the transgression of social norms. Under the veil of amusement, moral violation in self and other beneficial scenarios increased in a more important way when participants were exposed to dark humor. To deepen our understanding of the underlying mechanisms of the permissiveness within different humorous contexts, four main lines of research must be considered.

First, it's noteworthy that people are motivated by their self-interest and prone to behave for their own benefit in moral dilemmas (Christensen et al., 2014; Moore, Clark, & Kane, 2008; Moore, Stevens, & Conway, 2011). In our study, individuals tended to exhibit a utilitarian response style when the transgressive act was described as self and other beneficial as compared to other-beneficial. This effect is coherent with Greene's dual-process model (Greene et al., 2001, 2004, 2008), which suggests that utilitarian judgments result from a deliberate analysis of costs and benefits. In this cost-benefit perspective of moral judgment, saving oneself (in addition to others) could be considered as an additional benefit: The gains

represented by saving oneself and others outweigh the gains of saving others only. Hence, people probably experience less conflict in the analysis of cost-benefit ratio when action is for their own benefit too. This interpretation is compatible with Moore et al. (2008, 2011) who showed that individuals were faster to approve the “utilitarian” response when those who benefited from the crime included themselves. This interpretation is also consistent with Shalvi, Gino, Barkan, & Ayal (2015; see also Gino, Ayal, & Ariely, 2013) who underlined that people experience less internal conflicts when the temptation to profit from unethical behavior can be justified by saving others. This self-serving altruism could explain our results. In the present study, the self and other beneficial scenario enabled people to violate moral rules (e.g., approve a behavior that cause harm to a victim) while maintaining their positive self-image, because the moral violation also benefitted others.

Secondly, our results showed that the tendency to accept moral violation in both their self and other interest increased when participants were exposed to dark jokes. According to the benign violations theory literature (McGraw & Warren, 2010; Warren & McGraw, 2015, 2016), this result suggests that humorous contexts affect moral judgment via appraisal processes. Indeed, the fact that the permissiveness of the moral violation increases in self and other beneficial dilemmas under dark humorous context suggests that this informational context promotes the appraisal of utilitarian response as a benign violation. In other words, when negative serious ideas (like death or suffering of the victims) are associated to farcical elements through dark jokes, the utilitarian response probably becomes more benign, especially when it benefits oneself in addition to others. However, such explanation should be considered with caution since we did not directly emphasize the moral component of the utilitarian response. In the present study, the harmful actions described in sacrificial dilemmas were judged appropriate or inappropriate only. Future studies should directly investigate if utilitarian judgment is perceived as more or less immoral according to the type of humor

(dark vs. nondark) and the type of dilemma (Self and Other-beneficial vs. Other-beneficial). Another limitation, inherent to almost all moral dilemma research is about the nature of the task and the corresponding measure. As highlighted by Crone and Laham (2017), sacrificial dilemma responses have to be considered with caution since they have been proved to be a poor indicator of moral values. Obviously, there is a huge difference between what one judges as morally acceptable and what one actually does (see also Tassy, Oullier, Mancini, & Wicker, 2013). A profitable line of research would consist of distinguishing between evaluative judgments and choices of action.

Thirdly, the way the scenarios are perceived is another line of research that deserves to be considered. Bauman, McGraw, Bartels and Warren (2014) strongly recommend that researchers be cautious when using sacrificial dilemmas to studying moral judgment: The ecological validity of sacrificial dilemmas needs to be carefully considered (see Körner, Joffe, & Deutsch, 2019) because the lack of realism may threaten the validity of moral decision processes under interest. Also, because the scenarios are hypothetical, a utilitarian response (i.e., kill someone) could be seen as both a violation and benign. In this circumstance, the benign-violation predicts that people will be amused. In accordance with this hypothesis, Bauman et al. (2014) showed that, in the footbridge scenario, people considered pushing the man to be wrong, but also reported laughing. If sacrificial dilemmas have the power to elicit humor, we can hypothesize that exposure to dark jokes promotes the violation and benign appraisals of the situation described in the scenario, and thus, generates laughter. The question is can dark humor transform a serious scenario into a humorous one, because its transgressive content leads to perceive that moral violation is actually okay. Rather than abandoning sacrificial dilemmas entirely, using a virtual reality paradigm may offer a more vivid experience of the scenarios, making their realism more salient (e.g., McDonald, Defever, & Navarrete, 2017; Slater & Sanchez-Vives, 2016), and elicit more serious moral deliberation.

A last but not least contribution of the present study concerns individuals' emotional state when faced with moral dilemma. Contrary to previous studies that only considered nondark humor, our comparison between dark and nondark humor revealed that variability in moral judgment could not simply be explained in terms of experiencing positive emotions. Interestingly, the benign-violation theory (McGraw & Warren, 2010) suggests that humor does not systematically involve positive emotions (e.g., amusement, mirth). Because humor results from violations that are simultaneously seen as benign, it may elicit mixed emotions. This idea is in line with theorists (e.g., Larsen & McGraw, 2014; Larsen, McGraw, & Cacioppo, 2001; Schimmack, 2001) who argue that positive and negative emotions do not mutually inhibit each other, and may at times even co-occur (i.e., mix). Clearly, humor elicits mixed feelings such as disgust and amusement. For instance, people are both amused and disgusted when exposed to a disgusting humorous clip (Hemenover & Schimmack, 2007). Aillaud and Piolat (2014) provided additional evidence when underlying that participants used both positive and negative emotional lexicon to describe the emotional experience associated with dark and nondark humorous cartoons. These authors reported that dark humor is particularly conducive to provoking mixed emotions. Not only does its transgressive content elicit amusement, but also triggers negative emotions such as shame or disgust. Since the present study operationalized dark humor, participants may have felt amusement and shame (or/and disgust), two emotions of opposite valence. This hypothesis cannot be tested in our study since we measured emotional valence only. Our results revealed that participants reported feeling happier after reading the jokes than before reading the jokes, but their induced happiness decreased after the moral judgment task. Further research should consider the panel of discrete emotions to understand how individuals manage mixed feelings when asked to judge whether the acts are morally appropriate or not. It would be interesting to examine what they feel in the different steps: before the moral judgment task, during the

reading of the scenario and after the moral judgment. Because the dilemmas have proven to elicit different negative emotions (i.e., guilt, disgust, sadness, empathy, anger; see Choe & Min, 2011), the question remains on how different types of humor can counterbalance such negative feelings.

Overall, there is no doubt that the next step to overcome in deepening our understanding of moral judgments is to focus on its context-dependent nature. This line of research allows us to get a better understanding of the mechanisms in which humor influences moral judgment. Some additional factors may contribute to this line of research like an individuals' need for humor (Cline, Altsech, & Kellaris, 2003; see also Picard & Blanc, 2013) and also gender (e.g., Herzog & Anderson, 2000). Interestingly, in the present research, dark humor effects are observed on a sample composed exclusively of females. It is noteworthy that females are known to usually find less humor in dark events than males (Aillaud & Piolat, 2012). The tendency to produce a utilitarian response could be strongly reinforced under dark humor with males who are predisposed to generate and seek out humor (i.e., who scored high in need for humor). Future research is needed to test this hypothesis. Finally, this study sheds light on the necessity to not neglect the fact that moral judgments take place in a specific sociocultural environment more or less prompt to accept dark humor. The exposure to dark humor in an individualist culture is of great importance since moral decision experienced in everyday life is often driven by selfish attitudes. The presence of dark humor can promote moral transgression that favors the tolerance of utilitarian response.

5. References

Aillaud, M., & Piolat, A. (2012). Influence of gender on judgment on dark and nondark humor. *Individual Differences Research*, 10, 211-222.

- Aillaud, M., & Piolat, A. (2014). Nature du lexique émotionnel produit en situation d'appréciation et de rejet de dessins d'humour noir et d'humour non noir. *Canadian Journal of Experimental Psychology*, 68, 166-178.
- Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: The 'other-praising' emotions of elevation, gratitude, and admiration. *Journal of Positive Psychology*, 4, 105-127.
- Backs, R. W., da Silva, S. P., & Han, K. (2005). A comparison of younger and older adults' self-assessment manikin ratings of affective pictures. *Experimental Aging Research*, 31, 421-440.
- Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, 108, 381-417.
- Bartels, D. M., Bauman, C. W., Cushman, F. A., Pizarro, D. A., & McGraw, A. P. (2015). Moral judgment and decision making. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell handbook of judgment and decision making* (pp. 1-51). Chichester, UK: Wiley.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121, 154-161.
- Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, 8/9, 536-554.
- Bynion, T. M., & Feldner, M. T. (2017). Self-assessment Manikin. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences*. Cham, Switzerland: Springer.

- Charaudeau, P. (2006). Des catégories pour l'humour ? *Questions de communication*, 10, 19-41.
- Choe, S. Y., & Min, K. H. (2011). Who makes utilitarian judgments? The influences of emotions on utilitarian judgments. *Judgment and Decision Making*, 6, 580-592.
- Christensen, J. F., Flexas, A., Calabrese, M., Gut, N. K., & Gomila, A. (2014). Moral judgment reloaded: a moral dilemma validation study. *Frontiers in Psychology*, 5, 607.
- Christopher, S. (2015). An introduction to black humour as a coping mechanism for student paramedics. *Journal of Paramedic Practice*, 7, 610-615.
- Ciaramelli, E., Muccioli, M., Ladavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, 2, 84-92.
- Cline, T. W., Altsech, M. B., & Kellaris, J. J. (2003). When does humor enhance or inhibit ad responses? The moderating role of the need for humor. *Journal of Advertising*, 32, 31-45.
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, 104, 216-235.
- Crone, D. L., & Laham, S. M. (2017). Utilitarian preferences or action preferences? De-confounding action and moral code in sacrificial dilemmas. *Personality and Individual Differences*, 104, 476-481.

- Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*, 1-35.
- Ford, T. E., Boxer, C. F., Armstrong, J., & Edel, J. R. (2008). More than “just a joke”: The prejudice-releasing function of sexist humor. *Personality and Social Psychology Bulletin*, 34, 159-170.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56, 218-226.
- Freud, S. (1905). *Jokes and their relation to the unconscious*. Norton : New York.
- Gervais, M., & Wilson, D. S. (2005). The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly Review of Biology*, 80, 395-430.
- Gino, F., Ayal, S., & Ariely, D. (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of Economic Behavior and Organization*, 93, 285-292.
- Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PloS ONE*, 8, e60418.
- Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in cognitive sciences*, 11, 322-323.
- Greene, J. D. (2008). The secret joke of Kant’s soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 3: The neuroscience of morality: Emotion, brain disorders, and development* (pp. 35-117). Cambridge, MA : MIT Press.

- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144-1154.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*, 389-400.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105-2108.
- Gutiérrez, C. A., Carretero-Dios, H., Willis, G. B., & Morales, M. C. (2018). “It’s funny if the group says so”: Group norms moderate disparaging humor appreciation. *Humor: International Journal of Humor Research*, *31*, 473-490.
- Haidt, J. D. (2003). Elevation and the positive psychology of morality. In C. L. M. Keyes & J. D. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived* (pp. 275-289). Washington, DC: American Psychological Association.
- Hemenover, S. H., & Schimmack, U. (2007). That’s disgusting! ..., but very amusing: Mixed feelings of amusement and disgust. *Cognition & Emotion*, *21*, 1102-1113.
- Herzog, T. R., & Anderson, M. R. (2000). Joke cruelty, emotional responsiveness, and joke appreciation. *Humor: International Journal of Humor Research*, *13*, 333-351.
- Kahane, G., Everett, J. A. C., Earp, B. D., Farias, M., & Savulescu, J. (2015). ‘Utilitarian’ judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193-209.

- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, 446, 908-911.
- Körner, A., Joffe, S., & Deutsch, R. (2019). When skeptical, stick with norm: Low dilemma plausibility increases deontological moral judgments. *Journal of Experimental Social Psychology*, 84, 1-7.
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. William (Eds.), *Technology in mental health care delivery systems* (pp. 119-137). Norwood, NJ: Ablex.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings*. Gainsville, FL: NIMH Center for the Study of Emotion and Attention.
- Larsen, J. T., & McGraw, A. P. (2014). The case of mixed emotions. *Social and Personality Psychology Compass* 8/6, 263-274.
- Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, 81, 684-696.
- Martin, R., & Ford, T. (2018). *The psychology of humor: An integrative approach* (2nd ed.). Burlington: Elsevier Academic Press.
- McDonald, M. M., Defever, A. M., & Navarrete, C. D. (2017). Killing for the greater good: action aversion and the emotional inhibition of harm in moral dilemmas. *Evolution and Human Behavior*, 38, 770-778.

- McGraw, A. P., & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological Science, 21*, 1141-1149.
- McGraw, A. P., Warren, C., Williams, L., & Leonard, B. (2012). Too close for comfort, or too far to care? Finding humor in distant tragedies and close mishaps. *Psychological Science, 23*, 1215-23.
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science, 19*, 549-557.
- Moore, A. B., Stevens, J., & Conway, A. R. A. (2011). Individual differences in sensitivity to reward and punishment predict moral judgment. *Personality and Individual Differences, 50*, 621-625.
- Moretto, G., Ladavas, E., Mattioli, F., & di Pellegrino, G. (2009). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *Journal of Cognitive Neuroscience, 22*, 1888-1899.
- Morris, J. D. (1995). Observations: SAM: The Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. *Journal of Advertising Research, 35*, 63-68.
- Nabizadeh Chianeh, G., Vahedi, S., Rostami, M., & Nazari, M. A. (2012). Validity and Reliability of Self-Assessment Manikin. *Journal of Research in Psychological Health, 6*, 52-61.
- O'Connor, E. C., Ford, T. E., & Banos, N. C. (2017). Restoring threatened masculinity: The appeal of sexist and anti-gay humor. *Sex Roles, 77*, 567-580.

- Picard, D., & Blanc, N. (2013). Need for humor scale: Validation with French children. *Psychological Reports: Measures and Statistics*, 112, 502-518.
- Ramachandran, V. S. (1998). The neurology and evolution of humor, laughter, and smiling: The false alarm theory. *Medical Hypotheses*, 51, 351-354.
- Rowe, A., & Regehr, C. (2010). Whatever gets you through today: An examination of cynical humor among emergency service professionals. *Journal of Loss and Trauma*, 15, 448-464.
- Saucier, D. A., O'Dea, C. J., & Strain, M. L. (2016). The bad, the good, the misunderstood: The social effects of racial humor. *Translational Issues in Psychological Science*, 2, 75-85.
- Schimmack, U. (2001). Pleasure, displeasure, and mixed feelings: Are semantic opposites mutually exclusive? *Cognition and Emotion*, 15, 81-97.
- Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong, and feeling moral. *Current Directions in Psychological Science*, 24, 125-130.
- Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3, 74.
- Strohming, N., Lewis, R. L., & Meyer, D. E. (2011). Divergent effects of different positive emotions on moral judgment. *Cognition*, 119, 295-300.
- Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology*, 4, 250.
- Thai, M., Borgella, A. M., & Sanchez, M. S. (2019). It's only funny if we say it: Disparagement humor is better received if it originates from a member of the group being disparaged. *Journal of Experimental Social Psychology*, 85, 1-10.

- Thomson, J. (1985). The Trolley Problem. *Yale Law Journal*, 94,1395-1415.
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgement. *Psychological Science*, 17, 476-477.
- Van Wormer, K., & Boes, M. (1997). Humor in the emergency room: A social work perspective. *Health and Social Work*, 22, 87-92.
- Veatch, T. C. (1998). A theory of humor. *Humor: International Journal of Humor Research*, 11, 161-215.
- Waldmann, M. R., Nagel, J., & Wiegmann, A. (2012). Moral judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 364-389). New York: Oxford University Press.
- Warren, C., & McGraw, A. P. (2015). Opinion: What makes things humorous. *PNAS Proceedings of the National Academy of Sciences*, 112, 7105 -7106.
- Warren, C., & McGraw, A. P. (2016). Differentiating what is humorous from what is not. *Journal of Personality and Social Psychology*, 110, 407- 430.
- Woodzicka, J. A., & Ford, T. (2010). A Framework for Thinking about the (not-so-funny). Effects of Sexist Humor. *Europe's Journal of Psychology*, 3,174-195.
- Young, M. (1995). Black humor: Making light of death. *Policing and Society*, 5, 151-167.