



# Kent Academic Repository

**Gutierrez, Roberto and Giner-Sorolla, Roger (2011) *Disgusting but harmless moral violations are perceived as harmful due to the negative emotions they elicit*. *Revista de Psicología Social*, 26 (1). pp. 141-148. ISSN 0213-4748.**

## Downloaded from

<https://kar.kent.ac.uk/36667/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.1174/021347411794078381>

## This document version

Author's Accepted Manuscript

## DOI for this version

## Licence for this version

UNSPECIFIED

## Additional information

Translated title: Las transgresiones morales repulsivas pero inofensivas se perciben como dañinas debido a las emociones negativas que inducen

## Versions of research works

### Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Disgusting but harmless moral violations are perceived as harmful due to the negative emotions they elicit

Roberto Gutierrez

Anglia Ruskin University, United Kingdom

Roger Giner-Sorolla

University of Kent, United Kingdom

English translation (by the article's authors) of Spanish language article, published in *Revista de Psicología Social* (2011), 26, 141-148.

We thank Pablo Briñol and an anonymous reviewer for their comments on an early version of this manuscript.

Correspondence should be addressed to Roberto Gutierrez, Department of Psychology, Anglia Ruskin University, Cambridge, CB1 1PT. United Kingdom. E-mail: Roberto.Gutierrez@anglia.ac.uk

### **Abstract**

Harmless but disgusting moral violations can be justified as harmful to others due to the negative emotions they elicit. The relationship between the emotions of anger and disgust and the harm associated to these emotions as a result of a moral violation was investigated. Results showed that a disgusting moral violation (taboo violation) described as harmless to others is more related to disgust than to anger. Such violation created a presumption of harm of three different types: to the community, nature, and the individual. Disgust was a mediator between the taboo violation and the presumption of harm to nature, whereas anger was a mediator between the taboo violation and the presumption of harm to the individual. In general, results also showed that in moral violations that are harmless to others, the emotions of anger and disgust allow people to presume harm to symbolic entities such as nature and the community as a result of such violations.

**Disgusting but harmless moral violations are perceived as harmful due to the negative emotions they elicit**

The relationship between emotional reactions, moral violations and the evaluation of such action has experienced an important change in recent years. Until recently, most research focused on morality gave its attention on the rational processes that were presumed to be controlled and conscious (e.g., Kohlberg, 1971), leaving emotions and intuitions to play a secondary role at best (Haidt, 2001). More contemporary research investigating morality and moral judgement generally endorse that emotions and intuitions are a crucial component in the process of moral judgement and in the research of morality as a whole (Haidt, 2003). Some theoretical models now propose close relationships between the nature of the moral violations and specific emotions. For example, the CAD hypothesis (Rozin, Lowery, Imada, & Haidt, 1999) suggests a correspondence between a triad of moral violation domains (Community, Autonomy and Divinity) and three specific emotions (Contempt, Anger and Disgust). This model proposes that an action that violates a moral norm of the code of divinity—related to food, purity and the sanctity of the body—should elicit mostly, if not uniquely, disgust.

Although it is now accepted that emotions play a role in our evaluations and judgements (Damasio, 1994; Keltner & Gross, 1999), it is not clear how different emotions have different effects on the way we judge moral violations. For example, anger and disgust are two emotions that are frequently mentioned as responses towards moral violations. Disgust has been portrayed in several different forms, from core disgust to sociomoral disgust (Simpson, Carter, Anthony, & Overton, 2006). These wide range of ‘functions’ of disgust means that it can be implicated in a large number of phenomena, from the protection against contaminated food (Rozin, Haidt, & McCauley, 1999), to a justification for murder (Kahan, 1998). Anger

has been mostly referred in moral psychology as a response to an unjustified insult to one or on behalf of someone else and a common response to unfairness (Haidt, 2003).

### Anger, Disgust and Presumption of Harm

The close relationship between the two ‘other-condemning’ emotions (Haidt, 2003), anger and disgust, as well as the wide range of elicitors associated to them present an important challenge in moral judgement, particularly when these moral violations do not have negative consequences to others such as in the case of consensual homosexual relationships. One of the consequences associated to harmless moral violations is the so called ‘moral dumbfounding effect’. This effect refers to the condemnation of moral violations even when the person is unable to provide reasons to evaluate the action in a negative manner. Generally people condemn violations to moral norms without negative consequences to others, even when they can provide reasons to condemn such actions (Haidt, Bjorklund, & Murphy, 2004). It has been reported that this effect may be an *intuition*, allowing people to quickly and effortlessly reach a moral judgement without the need or reasons. According to the social intuitionist model (Haidt, 2001). People reach a moral judgement based on fast intuitions, and then use reasons to justify the negative evaluation. More specifically on emotions, the presumption of harm model (Gutierrez & Giner-Sorolla, 2007) suggests that people ascribe harm to moral violations that elicit negative emotions—anger in particular—even when these actions have no negative consequences on other people. In such investigation it was revealed that anger was associated to the belief that a moral transgression harmed the rights of other people, although the action was described as private and secret.

In this investigation we proposed that such *presumed harm* can be extended to symbolic entities such as nature and the community in cases in which no one has been actually harmed,

making it compatible with the moral dumbfounding effect and the social intuitionist model. This extension allows investigating whether other emotions apart from anger are implicated in the presumption of harm to symbolic entities such as the rights of other people, or the traditions of the community. In addition, another theoretical account suggests that disgust ‘moralises’ evaluations associated to violations of the domain of divinity, so that the presence of disgust makes evaluations about divinity violations more severe (Horberg, Keltner, Oveis, & Cohen, 2009).

Most research (e.g. The CAD hypothesis) relates the emotion of anger to violations of the domain of autonomy, which is usually associated to harm to the individual. However, in cases in which there is no real harm to an individual the emotion of disgust seems to provide an avenue to ascribe harm to a harmless moral violation in the form of presumed harm. This situation brings an interesting question: What emotion, anger or disgust, is associated to harm in harmless taboo violations? On one hand, the large majority of theoretical accounts have proposed that anger is the result of harm. On the other hand results have shown that in moral violations, disgust is closely related to harm in the form of a ‘moralising’ agent (Horberg et al., 2009), a presumption of harm to symbolic entities such as other’s rights (Gutierrez & Giner-Sorolla, 2007), or as an intuitive response (Haidt, 2001).

We propose that in the case of harmless taboo violations, specific emotions will be associated to different types of harm. As in, we refer here to taboo violation as norms whose violation can be expected to provoke inflexible, disgust-related responses and, in particular, norms related to the body, food, and sexuality. The different types of harm are related to the ‘big three’ moral codes proposal (Shweder, Munch, Mahaptra, & Park, 1997), so that the moralisation of moral violations due to disgust (Horberg et al., 2009), can be explained based on presumption of harm to nature; whereas the relationship autonomy violations and anger can be explained as presumption of harm to an individual.

It is expected that when participants are presented with a harmless moral violation, their evaluation will be negative based on a presumption of harm to a symbolic entity such as rights and the community. Also, it is expected that the reported harm can be separated in different types such as individual harm (violation of the autonomy code), community harm (violation of the community code), and harm to nature (violation of the divinity code). Finally, it is expected that the emotion of anger, but not disgust, will be most associated to individual harm; and that disgust, but not anger, will be associated to harm to nature.

## **Method**

### *Participants*

Seventy-nine participants of the Norwich Arts Centre and the Norwich Environmental Centre responded individually to a questionnaire on a voluntary basis. Of these, 43 were males and 36 were females.

### *Design*

This experiment had a single factor with 2 levels between participants design (Condition: Taboo vs. Control).

### *Materials*

The questionnaire consisted in a booklet that contained one fictitious story that was modified to create two different conditions, whether the main character of the story performed either a taboo action (a scientist eating a portion of meat cloned from her own arm) or a non-taboo action (a scientist drinking drug that alters memory). These stories were taken from Gutierrez and Giner-Sorolla (2007, Experiment 2). In both cases, no explicit actual harm to others was described as a result of the action of the main character of the story.

*Harm to nature:* All responses were given in bipolar scales from 1 (Not at all) to 9 (extremely.) Two items measured the perceived harm to nature as a result of the scientist

action. “Do you think the action of the scientist caused any damage to the natural order of things?” and “Do you think the action of the scientist violated the laws of nature?”

*Community harm:* Two items measured the perceived harm to the community as a result of the scientist action. “Do you think the action of the scientist caused any damage to the community?” and “Do you think the action of the scientist violated the rights of anyone apart from her?”

*Individual harm:* Three items measured the perceived harm to other people. “Do you think the action of the scientist caused any physical harm to anyone?”, “Do you think the action of the scientist caused any psychological harm to anyone?” and “Do you think the action of the scientist caused any emotional harm to anyone?”

*Evaluation of the action:* The evaluation of the action of the main character of the story was measured with four items in the form of bipolar scales from 1 (completely right / good / correct / positive) to 9 (Completely wrong / bad / incorrect / negative).

*Emotion items.* In the *verbal emotion* items, participants were asked to indicate to what extent each story made them feel anger, compassion, depression, disgust, happiness, infuriation, outraged, pity, pleasure, repulsion, sadness, satisfaction, sickness, sorrow, sympathy, grossed-out and contempt. These measures were answered in scales from 1 (not at all) to 8 (very). In the *facial emotion* items, two photographs of female faces were shown, one showing disgust in the full form and the other showing anger in the open mouth form. Both expressions fulfil the requirements to be considered prototypical emotions based on the Facial Action Coding System (FACS, Ekman & Friesen, 1978). Participants were instructed to "select the face that best describes your feelings towards the scientist now". Immediately after selecting a facial expression, participants were asked to indicate separately how much of each of the feelings represented by each face they had towards the scientist in scales from 1 (not at all) to 9 (extremely).



*Procedure*

After the presentation of the story the participants were requested to answer the items regarding the harm to nature, the community and the individual. The order of the questions was changed in a cyclical way, so that some participants read the questions related to harm to nature first, other participants read the questions associated to the community harm first, and other participants read the harm to nature questions first. After that, all participants answered the items regarding the evaluation of the action, followed by the verbal and facial measures of emotions. Finally, all participants were thanked and debriefed.

**Results***Evaluation of the action*

The four evaluation items (Cronbach's alpha = .88) were averaged to create one single score. Analysis of variance on the evaluation score showed a significant main effect of Condition  $F(1, 79) = 8.71$ ,  $MSE = 2.60$ ,  $p < .01$ , indicating that the taboo action was evaluated more negatively than the non-taboo action. Means and standard deviations are shown in

Table 1.

Table 1. *Mean and standard deviation of evaluations, harm and emotions by condition*

	Condition	
	No Taboo	Taboo
Evaluation	4.56 (1.24)	5.63 (1.94)
Individual harm	2.56 (1.58)	3.53 (2.03)
Harm to nature	3.01 (2.18)	5.38 (2.89)
Harm to community	2.46 (2.03)	3.13 (2.16)
Verbal anger	1.70 (1.69)	2.18 (1.28)
Verbal disgust	1.51 (.82)	3.39 (2.19)
Facial anger	1.68 (1.21)	2.29 (2.04)
Facial disgust	1.95 (1.76)	3.61 (2.65)
Combined score Anger	-.18 (.74)	.19 (.90)
Combined score Disgust	-.41 (.45)	.44 (1.05)

*Note:* Standard deviations are in parenthesis.

#### *Individual, Nature and Community Harm*

Principal components analysis with varimax rotation was used to explore the existence of the proposed 3 types of harm. A factor analysis requesting 3 factors explained 83.42 % of the variance and showed three distinctive factors, one related to harm to the individual, one related to harm to nature and one related to harm to the community. Confirmatory factor analysis was used to compare the fitness of the data with three different models. A two-factor model (Individual harm vs. nature-community harm), a three-factor model (individual vs. nature vs. community) in which the factors were not correlated, and a three-factor model (individual vs. nature vs. community) in which the factors were correlated. Results revealed

that the third model was better than the other two, being the one showing the best fit (Table 2).

A mixed model analysis of variance with a 3 (Harm: Natural vs. Community vs. Individual, within participants factor) x 2 (Condition: Taboo vs. No Taboo, between participants factor) revealed significant main effects of Harm,  $F(2, 154) = 16.77$ ,  $MSE = 2.62$ ,  $p$

$< .001$ , and Condition  $F(1, 77) = 11.85$ ,  $MSE = 8.92$ ,  $p < .001$ . Importantly, these main effects were qualified by a significant Harm x Condition interaction  $F(2, 154) = 6.20$ ,  $MSE = 2.62$ ,  $p < .01$ . Simple effect analysis revealed that the taboo manipulation had no significant effect on harm to the community  $F(1, 77) = 2.00$ ,  $MSE = 4.40$ ,  $p = .16$ , whereas there was a significant effect on individual harm,  $F(1, 77) = 5.67$ ,  $MSE = 3.28$ ,  $p < .05$ , increasing the perception of harm in the Taboo condition compared to the No Taboo condition; and an even larger effect of harm to nature  $F(1, 77) = 17.09$ ,  $MSE = 6.48$ ,  $p < .001$ .

Table 2. *Confirmatory factor analysis results and goodness of fit indicators*

	$\chi^2$	df	p	RMSEA	CFI	NFI
Two factors	44.74	13	<.001	.19	.89	.86
Three uncorrelated factors	64.80	16	<.001	.20	.83	.79
Three correlated factors	6.74	11	.82	0	1.00	.98

*Emotions*

*Verbal measures.* The 3 items related to anger (anger, infuriation and outraged; Cronbach's alpha = .78), and the 4 items regarding disgust (disgust, sick, repulsed and grossed-out; Cronbach's alpha = .93) were averaged to create one index for each emotion. These indexes were correlated at  $r(77) = .61, p < .001$ . A mixed model analysis of variance adding a 2 (Emotion: Anger vs. Disgust, within participants) factor to the basic design revealed significant main effects of Emotion  $F(1, 77) = 12.02, \text{MSE} = 0.87, p < .001$ , and Condition  $F(1, 77) = 16.73, \text{MSE} = 3.29, p < .001$ . A significant interaction between these factors was also present  $F(1, 77) = 23.35, \text{MSE} = 0.87, p < .001$ . Simple effect analysis revealed that the manipulation had a stronger effect on the index of disgust,  $F(1, 77) = 26.29, \text{MSE} = 2.66, p < .001$ , than on the index of anger,  $F(1, 77) = 3.04, \text{MSE} = 1.50, p = .09$ . Means and standard deviations are in

Table 1.

*Facial measures.* Nine participants did not select any of the faces, all of them in the control condition. Analysis of the remaining responses revealed that the correlation between the faces selected was  $r(77) = .41, p < .001$ . The face representing disgust was selected more overall (53 times, 75.7 %), than the face representing anger (17 times, 24.3 %). Further analysis revealed that this pattern was repeated for both conditions, but the effect was stronger in the Taboo condition.

The face representing anger was more correlated with the index related to anger words  $r(77) = .41, p < .001$ , than with the index representing disgust words,  $r(77) = .23, p < .05$ , this difference was marginally significant  $t(76) = 1.94, p = .06$ . Likewise, the face representing disgust was more correlated with the disgust index  $r(77) = .63, p < .001$ , than with the anger index  $r(77) = .39, p < .001$ , this difference was significant  $t(76) = 2.97, p < .01$ . A mixed model analysis of variance with a 2 (Face: anger vs. disgust within participants) factor, replacing the words measures with the facial measures revealed significant main effects of Face  $F(1, 77) = 9.97, \text{MSE} = 2.48, p < .01$ , and Condition  $F(1, 77) = 9.59, \text{MSE} = 5.26, p < .01$ , as well as a significant interaction between these factors  $F(1, 77) = 4.36, \text{MSE} = 2.48, p < .05$ , indicating that there was a significantly higher reported level of disgust than anger in the Taboo condition,  $t(37) = 2.84, p < .01$ , but no significant difference between the emotions in the No Taboo condition  $t(40) = 1.21, p = .23$  (

Table 1).

One index for each emotion was created averaging the standard value of the faces and words measures of each emotion. These scores were correlated at  $r(77) = .54, p < .001$ . Analysis of such scores using a 2 (Emotion: anger vs. disgust, within participants factor) x 2 (Condition: Taboo vs. No Taboo, between participants factor) mixed model analysis of variance revealed a significant main effect of Condition,  $F(1, 77) = 14.94, \text{MSE} = 0.99, p < .001$ , suggesting a higher overall indication of anger and disgust in the disgust condition than in the control condition. The main effect of Emotion was not significant,  $F(1, 77) = .01, \text{MSE} = .019, p = .92$ , but it was qualified by a significant Emotion x Condition interaction,  $F(1, 77) = 6.91, \text{MSE} = 0.32, p < .05$ . These results suggest that the Taboo action had a significant effect on both emotions, but a stronger effect on disgust,  $F(1, 77) = 22.36, \text{MSE} = 0.64, p < .001$ , than on anger,  $F(1, 77) = 4.09, \text{MSE} = 0.68, p < .05$ .

#### *Type of Harm and Emotions*

The relationship between the 3 types of harm and the emotions of anger and disgust was analysed using a bootstrap technique following the procedure and macros in Preacher and Hayes (2004). The independent effect of the Taboo manipulation was entered as a predictor of the 3 types of harm in separate analysis and the 2 combined indexes of anger and disgust were entered as simultaneous mediators between the taboo manipulation and the type of harm analysed. Results using unstandardised coefficients (B values) showed that the manipulation had a significant effect on both emotions, but it was stronger on disgust than in anger. In addition, the manipulation was significant on individual harm and harm to nature, but not on the community harm. As expected, anger but not disgust significantly predicted individual harm (Figure 1), whereas disgust but not anger significantly predicted harm to nature (Figure 2). Harm to the community was not predicted by anger or disgust (Figure 3).

Figure 1. *Anger, disgust and taboo manipulation on individual harm*

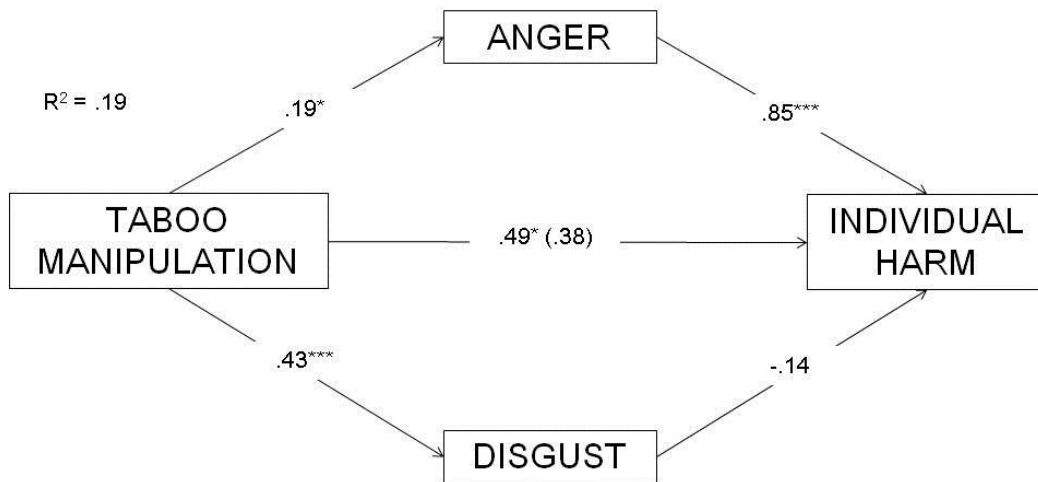




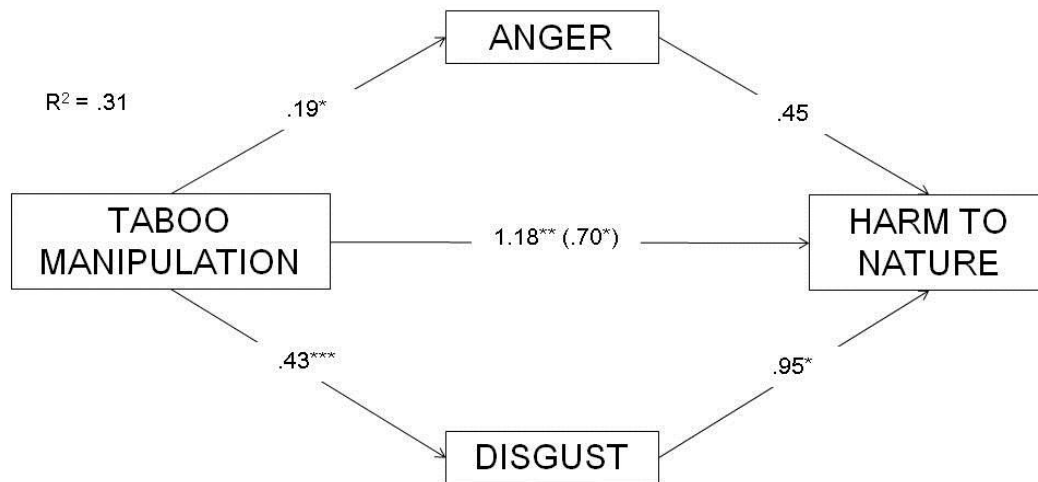
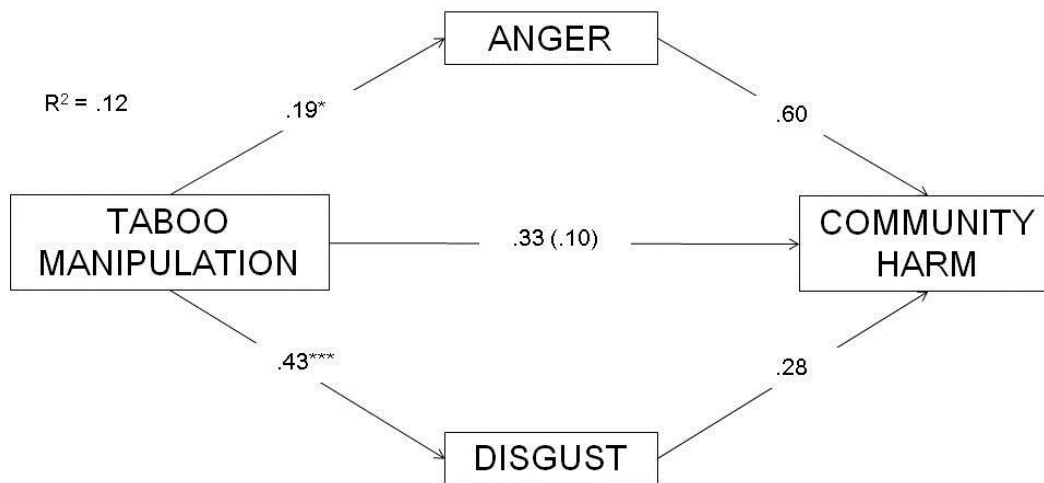
Figure 2. *Anger, disgust and taboo manipulation on harm to nature*

Figure 3. *Anger, disgust and taboo manipulation on community harm*

The taboo manipulation had a clear effect on the evaluation of the action and on the emotions of anger and disgust. It was expected that such manipulation would make the moral judgement of the action more severe even when no negative consequences were described in the scenarios.

The results confirmed that participants not only evaluated the taboo actions more negatively, but they associate harm to them. The presence of harm even when is none described is consistent with the presumption of harm model (Gutierrez & Giner-Sorolla, 2007), in which participants ascribe harm to symbolic entities in order to justify negative emotions and evaluations to a harmless but disgusting act.

### **Discussion**

These results also provide support for the ‘big three’ theory of morality (Shweder et al., 1997), separating moral violations in the domain of autonomy, community and divinity. These results also provide support to the CAD hypothesis, which proposes a correspondence between moral violations and specific emotions (Rozin et al., 1999). The results in this experiment showed that the taboo manipulation—a violation of the divinity code—had a stronger effect on the emotion of disgust than anger, and a stronger effect on the natural harm than the other types of harm. Importantly, the results presented here suggest that a moral violation is not restricted to one moral code, but that a violation of the code of divinity also has an effect on the code of autonomy and the emotion of anger.

These results suggest that private actions without negative consequences to others may give rise to the presumption of harm from negative emotions such as anger and disgust. It is important to note that this experiment described an action related to “core disgust”, which is related to contamination and food. There is a possibility that other types of moral violations that cause other types of disgust can be associated more to harm to the community. An

example is the case of homosexuality, which in some cases has been associated with socio-moral disgust, arguing that the existence of homosexual relationships is a threat to the institution of the family.

### References

Damasio, A. R. (1994). *Descartes's error: Emotion, reason and the human brain*. London: Papermac.

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System*. Palo Alto, California: Consulting Psychology Press Inc.

Gutierrez, R., & Giner-Sorolla, R. (2007). Anger, disgust, and presumption of harm as reactions to taboo-breaking behaviors. *Emotion*, 7(4), 853–868.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.

Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852-870). Oxford: Oxford University Press.

Haidt, J., Bjorklund, F., & Murphy, S. (2004). Moral dumbfounding: When intuition finds no reason. University of Virginia.

Horberg, E., Keltner, D., Oveis, C., & Cohen, A. (2009). Disgust and the Moralization of Purity. *Journal of Personality and Social Psychology*, 97(6), 963–976.

Kahan, D. M. (1998). The Anatomy of Disgust in Criminal Law. *Michigan Law Review*, 96(6), 1621-1657.

Keltner, D., & Gross, J. J. (1999). Functional account of emotions. *Cognition and Emotion*, 13(5), 467-480.

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive development and epistemology* (pp. 151-235). New York: Academic Press.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect

effects in simple mediation models. *Behavior Research Methods*, 36(4), 717.

Rozin, P., Haidt, J., & McCauley, C. R. (1999). Disgust: The body and soul emotion. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 429-445). John Wiley & Sons Ltd.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76, 574-586.

Shweder, R. A., Munch, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In P. Rozin & A. Brandt (Eds.), *Morality and health* (pp. 119-169). New York: Routledge.

Simpson, J., Carter, S., Anthony, S. H., & Overton, P. G. (2006). Is disgust a homogeneous emotion? *Motivation and emotion*, 30(1), 31-41.