Journal of Articles in Support of the Null Hypothesis
Vol. 20, No. 1
 1539–8714
www.jasnh.com

The null effect of recalling an experience to elicit disgust: A replication and extension of Sato and Sugiura (2014)

Akiko Matsuo
Research Center for Advanced Science and Technology, The University of Tokyo

Tomoya Mukai
Graduate Schools for Politics and Law, The University of Tokyo

Yuri Taknaka
Tama University, School of Management and Information Sciences

The amplification hypothesis of disgust, which proposes that incidental disgust causes harsher moral judgment, remains controversial due to mixed results in previous studies. However, some studies have investigated further effects based on the hypothesis. This may be problematic, because it is not guaranteed that an amplification effect will occur first. Therefore, the present study tested whether the amplification hypothesis is itself supported by replicating a previous study conducted and published in Japan. This study was conducted as an extension of the original study's methodology, with a few modifications due to the COVID-19 pandemic situation. Our results showed small, non-significant effects. This study offers a valuable contribution to introducing replication to international readership.

Keywords: disgust, moral judgment, amplification hypothesis, replication

Corresponding Author: Akiko Matsuo, Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8904, Japan. Email: matsuo.akiko.1005@gmail.com

Acknowledgement: We would like to thank Dr. Atsushi Sato, the first author of the original study, for providing the materials used in the original study.

Author contribution: All the authors contributed to the conception and design of this study. AM and YT collected the data. TM processed the data and performed the overall implementation. All authors contributed to the interpretation of the results and the writing, reading, editing, reviewing, and approval of the manuscript.

Ethical statement: All procedures were approved by the Research Ethics Committee of Tama University. We explained the purpose of the study to all participants and obtained informed consent before participation. The participants were informed that they were free to drop out of the study at any time without any adverse consequences.

Introduction

Research on the role of emotions in morality has accumulated over the past 30 years. In particular, researchers have focused on disgust and its effects on moral judgment because of the unique phenomena triggered by disgust and the functions that disgust may have. Disgust was found to make the subsequent moral judgment more severe, even though that feeling is not related to moral judgment (i.e., incidental). For example, in a study by Eskine et al. (2011), participants who drank a bitter drink that drew gustatory disgust made harsher moral judgments than those who drank neutral and sweet drinks.

In a meta-analytic review, Landy and Goodwin (2015) distinguished three hypotheses that researchers were confused about, expressing three relational patterns between disgust and moral judgment: elicitation, amplification, and moralization. The elicitation hypothesis posits that moral transgressions evoke disgust. This hypothesis is not discussed further because this study focuses on the causal link between disgust and moral judgment. The amplification hypothesis states that disgust produces a more severe condemnation of moral transgressions, whereas the moralization hypothesis argues that disgust makes even non-moral actions morally condemned. Because research on the moralization hypothesis is scarce and it is difficult to obtain conclusive effect sizes, the present study mainly addresses the amplification hypothesis.

Landy and Goodwin (2015) clarified these different theories and meta-analyzed previous research, up to 2014, on the relationship between disgust and moral judgment. They observed a small, but reliable, amplification effect (d = 0.11). The inconsistent patterns of past findings can be partly attributed to the methodology used, which includes how disgust is elicited and what kind of moral judgment is made. The effect sizes obtained by different stimuli that elicit disgust vary from d = 0.01 to 0.37 and those by different kinds of moral violations from d = 0.10 to 0.11 (Landy & Goodwin, 2015). After 2015, some replication attempts in related studies failed, making the amplification effect controversial (Ghelfi et al., 2020).

One problem resulting from this inconsistency in previous research is that other streams of research have been conducted based on the assumption that disgust leads to more severe moral judgments, such as the cleansing effect. The cleansing effect explains why cleansing behavior (e.g., handwashing) after heightened disgust alleviates moral judgment (Schnall et al., 2008). As Johnson et al. (2014) failed to replicate the original study, it is unclear whether disgust intensifies the wrongness of moral transgressions in the first place.

In addition, almost all prior studies have been conducted in Western contexts. Despite the recent need to study the effects of disgust in non-Western cultures, few studies have been conducted with Japanese participants (Kudo, 2019; Sato & Sugiura, 2014). Even then, these studies did not focus on the effect of disgust on moral judgment but mainly addressed different effects (i.e., how dispositional mindfulness helps ease harsh moral judgment by decreasing disgust and the cleansing effect). Therefore, it is crucial to directly test the amplification hypothesis regarding disgust and moral judgment in non-Western cultures. This study contributes not only to confirming and extending the amplification hypothesis but also to conducting future research in which it serves as the basic assumption. Due to the circumstances of the COVID-19 pandemic, this study employed disgust-inducing stimuli that could be conveniently presented online.

We replicate part of the study conducted originally by Sato and Sugiura (2014). This work makes a significant contribution to the extant literature on disgust by introducing an original study with an effect size of over 0.80, published in Japanese, to the international readership and examining its reproducibility. The original authors investigated the effect of individual differences in mindfulness on moral judgment intensified by disgust. Mindfulness is defined as paying attention to the present moment without making evaluative judgments. The original authors (Sato & Sugiura, 2014) hypothesized that individual differences in this perception strategy would prevent participants from the automatic influence of elicited disgust on subsequent moral judgments. Their findings indicated that the degree of moral judgment varied depending on the mindfulness subscale. Those who scored low on the "acting with awareness" subscale made harsher moral judgments compared to those who scored high, whereas their scores on other subscales did not influence their judgments. Since mindfulness was not of interest in the present study, we replicated the induction of disgust (vs. neutral) and its effect on moral judgment. Specifically, the present study was conducted online with a sample of participants over 18 years (i.e., not restricted to college students) and used a different neutral scenario to ask them to write about what they did at home. The original study collected data from college students using paper and pencil in a face-to-face setting and used a neutral scenario to ask them to write about what their class was like in a large classroom. It was predicted that writing about a disgusting experience would make the subsequent moral judgment harsher than writing about a neutral experience.

Methods

Participants

The original study (Sato & Sugiura, 2014) did not report the effect size, thus we calculated it using the reported data (Ns = 36, 42; Ms = 14.49, 19.33; SEs = 0.93, 0.84).[1] This calculation yielded d = 0.88, with a 95% confidence interval (95% CI), ranging from 0.41 to 1.34. Based on the lower limit of the 95% CI (0.41), a sample size of 440 was required. Therefore, to account for some attrition, we aimed for 500 participants. A total of 502 participants who took part in our online survey were recruited from Japan through the Internet crowdsourcing service CrowdWorks. Each participant earned 70 yen for their participation, and the survey was designed to appear to participants as two independent surveys. Among the

[1] The abovementioned meta-analysis reports the effect size of Sato and Sugiura (2014) to be d = 0.73. However, judging from the statistics written in the text, their reported value is likely to be mistaken, probably due to a linguistic limitation.

original sample of 502 participants, two responses were from the same IP address and the other two were unmatchable with the ID provided by CrowdWorks and were therefore excluded. Responses from the remaining 498 participants were used for data analysis (228 men, 267 women, and 3 others). Participants' ages ranged from 19 to 74 years (M = 41.25, SD = 10.81).

Materials

The first author was contacted to obtain the materials used in the original study. Scenarios for the disgusted and neutral conditions, moral judgment items, and mood check items were obtained. As noted previously, the scenario for the neutral condition was changed in the present study because the participants were not necessarily college students.

Moral judgment descriptions. The moral judgment items comprised 16 descriptions, of which 11 were moral-related and five were non-moral. Out of the 11, seven were "moral-personal" descriptions (i.e., more emotional), and four were "moral-impersonal" descriptions. Furthermore, three of the seven involved low-conflict moral-personal descriptions (i.e., less emotional) and the other four involved high-conflict moral-personal descriptions. Respondents rated the appropriateness of each description on a Likert scale ranging from 0 (no problem at all) to 9 (strongly wrong). The description examples include, "You are making brownies for yourself. Although the recipe needs a cup of walnuts, you like macadamia nuts rather than walnuts" (non-moral); "You are looking for a job. You can lie about your background on the resume" (moral-impersonal); "When you were driving, you found an injured person who asked for a ride to a nearby hospital. If you drive the person, the leather-covered backseat would be ruined" (low-conflict moral-personal); and "A trolley is barreling down a track that will kill five people unless diverted. You are standing on a bridge with a fat man next to you. If you push the fat man onto the track, the trolley will derail, sparing five people" (high-conflict moral-personal).

Mood-check items. To check if disgust is properly induced, participants were asked to rate their present mood with seven items: "relaxed," "angry," "happy," "sad," "fearful," "depressed," and "disgusted." Respondents rated the items on a Likert scale ranging from 1 (not feeling it at all) to 7 (strongly feeling it). These items were presented before the disgust or neutral induction task and after all moral judgment tasks. Administering mood-check items immediately after disgust and neutral induction could possibly lead participants to properly attribute their change in mood to the task and would not have an effect on moral judgment. Thus, the second set of mood check items were presented after the moral judgment tasks, following the procedure in the original study.

Procedure

Participants believed that they were participating in two independent studies at the time of recruitment. The front page of the online survey stated that the first part of the research was about human memory and the second was about one's impressions of people's various behaviors. After responding to the mood-check items as a baseline measurement, participants were randomly assigned to either the disgust or neutral condition. In the disgust condition, participants were asked to recall and type their own experiences in which they felt sick and were inclined to vomit when seeing or touching something disgusting. Some examples involved using a filthy bathroom and finding rotten, foul-smelling food. In the neutral condition, participants were asked to recall and type their experiences at home on the previous day. Examples included waking up in response to an alarm clock and brushing one's teeth. Participants worked on this induction task for six minutes and then rated the item on how vividly they recalled the experience on a Likert scale ranging from 1 (not at all vividly) to 7 (very vividly).

Subsequently, the message on the screen announced that the first study was completed and the second was starting. In the "second study," the participants were asked to judge how wrong the behavior in each of the 16 descriptions mentioned above was. The mood-check items were then administered again. Finally, participants were debriefed. In the original study (Sato & Sugiura, 2014), a positive mood was induced in participants in the disgust condition after debriefing, and mood-check items were administered again to verify whether their disgust disappeared. Because the present study was conducted online, pictures of animals were presented to eliminate disgust.

Anaytical Strategy

The main topic of the original study was the moderating role of mindfulness; thus, it was incorporated into the model and its effects were estimated. However, because the present study did not focus on mindfulness, its effects were excluded from the analytical models. Specifically, in the manipulation check, the original study conducted an analysis of covariance (ANCOVA) using the treatment and dichotomous variables measuring the degree of mindfulness as independent variables and mood of disgust, which was measured before the treatment, as a covariate; whereas, we conducted an ANCOVA using only the treatment as an independent variable. Moreover, in contrast to the original study that introduced the degree of mindfulness as an independent variable and other dimensions of mindfulness as covariates, we conducted a series of four Welch's t-tests to test for differences between the conditions.

Results

Manipulation check. The results of the ANCOVA revealed no detectable differences between the disgust (N = 249, M = 3.72, SD = 1.91) and control (N = 249, M = 3.87, SD = 1.92) conditions (F(1, 495) = 0.82, $\eta_p^2 < 0.01$, Hedges' d = 0.08 [–0.10, 0.25], p = 0.37) regarding the mood-check items measured at the end of the survey.

Effects on moral judgment. According to the moral intuitionist theory (Haidt & Bjorklund, 2008), moral intuition can affect actions without conscious awareness. Therefore, although the manipulation did not succeed in inducing (conscious) disgust among participants, it is still possible that the unconscious priming

Table 1. Effects of disgust manipulation on four categories of moral judgment.

Category	Condition				t	p	d	95%CI			
	Control		Disgust								
Personal, low-conflict	22.81	4.25	22.67	4.81	$t(488.47) = 0.36$	0.72	0.03	[	-0.14	, 0.21	]
Personal, high-conflict	24.65	7.14	25.01	7.12	$t(496.00) = -0.56$	0.58	-0.05	[	-0.23	, 0.13	]
Impersonal	26.57	5.50	26.39	5.85	$t(494.16) = 0.36$	0.72	0.03	[	-0.14	, 0.21	]
Non-moral	8.28	6.12	8.06	6.89	$t(489.18) = 0.38$	0.71	0.03	[	-0.14	, 0.21	]

of disgust affected participants' moral judgment. Thus, we proceeded with a series of four Welch's *t*-tests between conditions. However, the results revealed no detectable differences among the four moral judgment categories (Table 1).

Discussion

This study investigated the amplification effect of induced disgust on subsequent moral judgments. The results showed that disgust was not properly induced in the first place, which was also found by Kudo (2019) using a Japanese sample. There are several possible reasons for this failure. First, the study was conducted online and did not detect inattentive participants. Peyton et al. (2022) found that replication studies conducted online during the COVID-19 pandemic had lower effect sizes than the original studies conducted in the pre-COVID period, offering participants' inattention in web surveys as a partial explanation. The method of inducing disgust used in the original study (i.e., recalling one's own disgusting experience and writing it down) may not be suitable for online implementation. Second, the induction methodology itself may not have worked because of its small effect size. Landy and Goodwin (2015) showed $d = 0.04$ as the estimated effect size of the imagined disgust induction contained in the 95% CI of our results ($d = 0.08$ [-0.10, 0.25]). Disgust, as well as other discrete emotions (Lench et al., 2011), are induced more efficiently with visual presentations than with other types of induction. Therefore, induction strategies should be carefully selected. Finally, some unknown concepts may have been involved in inducing disgust, such as one's tendency towards multiple aspects of mindfulness presented in the original study. In Sato and Sugiura (2014), the degree of induced disgust depended on individual differences in the subscales of the Five Facet Mindfulness Questionnaire. Likewise, incidental disgust was found to be associated with religiosity (Ritter et al., 2016), need for affect (Sun, Yeo, McKasy, & Shugart, 2019), preference for structure (Donato & Miceli, 2020), level of construal (Moran et al., 2021), and sensitivity to bodily sensations (Schnall et al., 2008), among others. Thus, disgust induction may be susceptible to other variables, and the present study could not capture their effects.

Although the present study failed to manipulate this, it remains unclear whether an amplification effect occurs. As Johonson et al. (2014) argue, it is important to reflect on our interpretation of the replication results, which are inconsistent with the original studies. Compared to emotions, such as fear and anger, inducing disgust can be a complex issue (Sun et al., 2019). The categories of induced disgust may affect the following tasks differently (e.g., core, animal-reminder, pathogen, and moral; Olatunji et al., 2009). In the present study, the participants' descriptions varied with respect to the type of disgust. Furthermore, while some studies have demonstrated that disgust can be induced at an unconscious level (Lee et al., 2020), Białek et al. (2021) found that participants' (conscious) feelings of disgust affect moral judgment relative to the physical prevalence and degree of disgusting stimuli. Variations in the types and subjective "feelings" of disgust may explain the failure to induce disgust in this study.

Regarding effect sizes for the comparison of disgust vs. neutral conditions about moral judgment, the present study indicated a range of $d = -0.05$ and 0.03. The effect size in the original study ($d = 0.88$) appeared to be an outlier among the previous studies addressed in the meta-analysis. The extremely large effect size in the original study could be attributed to cultural differences in the meta-analyses. In the present study, the effect sizes were small and relatively similar to those reported previously. Therefore, our results rule out this possibility. One disgust-related theory argue that it may have evolutionary functions independent of culture (Rozin et al., 2009).

In conclusion, our results suggest substantially smaller effect sizes for the amplification effect of disgust than in the original study and fail to replicate its findings. Additionally, the present study failed to experimentally elicit disgust; however, this does not mean that an amplification effect would not be observed when other methodologies are employed. Rather, the present study implies that researchers cannot take successful manipulation for granted for possible reasons such as the mode of stimuli presentation and individual differences. The world is still affected by the COVID-19 pandemic; hence, future research should attempt to identify convenient and robust ways to induce disgust, and complete a step-by-step replication of the amplification effect.

References

Białek, M., Muda, R., Fugelsang, J., & Friedman, O. (2021). Disgust and moral judgment: Distinguishing between elicitors and feelings matters. *Social Psychological and Personality Science, 12*(3), 304–313.

Donato, C., & Miceli, G. (2020). The effect of disgust on preference for structure: Evidence for a double-sided response. *Journal of Consumer Behaviour, 19*(1), 68–79.

Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science, 22,* 295–299.

Ghelfi, E., Christopherson, C. D., Urry, H. L., Lenne, R. L., Legate, N.,

Ann Fischer, M., ... & Sullivan, D. (2020). Reexamining the effect of gustatory disgust on moral judgment: A multilab direct replication of Eskine, Kacinik, and Prinz (2011). *Advances in Methods and Practices in Psychological Science, 3*(1), 3–23.

Haidt, J., & Bjorklund, F. (2008). Social intuitionists answer six questions about morality. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality* (pp. 181–217). MIT Press.

Johnson, D. J., Cheung, F., & Donnellan, M. B. (2014). Does cleanliness influence moral judgments? A direct replication of Schnall, Benton, and Harvey (2008). *Social Psychology, 45*(3), 209.

Kudo, E. (2019, September 11–13). *Moral judgments and cleanliness: A replication of Schnall, Benton, & Harvey (2008)* [Conference presentation]. The 83rd Annual Convention of the Japanese Psychological Association, Ritsumeikan University, Osaka, Japan.

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science, 10*(4), 518–536.

Lee, K. M., Lindquist, K. A., Arbuckle, N. L., Mowrer, S. M., & Payne, B. K. (2020). An indirect measure of discrete emotions. *Emotion, 20*(4), 659.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137,* 834–855.

Moran, T., Bornstein, O., & Eyal, T. (2021). The level of construal involved in the elicitation of core versus moral disgust. *Emotion, 21*(2), 391–404.

Olatunji, B. O., Moretz, M. W., McKay, D., Bjorklund, F., de Jong, P. J., Haidt, J., ... & Schienle, A. (2009). Confirming the three-factor structure of the disgust scale—Revised in eight countries. *Journal of Cross-Cultural Psychology, 40,* 234–255.

Peyton, K., Huber, G. A., & Coppock, A. (2022). The generalizability of online experiments conducted during the COVID-19 pandemic. *Journal of Experimental Political Science, 9*(3), 379–394.

Ritter, R. S., Preston, J. L., Salomon, E., & Relihan-Johnson, D. (2016). Imagine no religion: Heretical disgust, anger and the symbolic purity of mind. *Cognition and Emotion, 30*(4), 778–796.

Rozin, P., Haidt, J., & Fincher, K. (2009). From oral to moral. *Science, 323*(5918), 1179–1180.

Sato, A., & Sugiura, Y. (2014). Dispositional mindfulness modulates automatic transference of disgust into moral judgment. *Japanese Journal of Psychology, 84,* 605–611.

Schnall, S., Benton, J., & Harvey, S. (2008). With a clean conscience cleanliness reduces the severity of moral judgments. *Psychological Science, 19,* 1219–1222.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096–1109.

Sun, Y., Yeo, S. K., McKasy, M., & Shugart, E. (2019). Disgust, need for affect, and responses to microbiome research. *Mass Communication and Society, 22*(4), 508–534.

Received: 4.17.2023
Revised: 6.7.2023
Accepted: 6.8.2023