

# Beyond Outrage: Observers Anticipate Different Behaviors From Expressors of Anger Versus Disgust

Lei Fan<sup>1</sup> , Catherine Molho<sup>1</sup> , Tom R. Kupfer<sup>2</sup>, Disa A. Sauter<sup>3</sup>, and Joshua M. Tybur<sup>1</sup> 

Social Psychological and  
Personality Science  
2024, Vol. 15(4) 450–460  
© The Author(s) 2023



Article reuse guidelines:  
zsagepub.com/journals-permissions  
DOI: 10.1177/19485506231176954  
journals.sagepub.com/home/spp



## Abstract

The modern world affords unprecedented opportunities for individuals to express moral sentiments. The widespread distribution of one specific type of sentiment — outrage — has consequences for social and political harmony. The current investigation contributes to better understanding these consequences by examining what types of aggression people expect from the outraged. Furthermore, it delineates how these expectations are shaped by the emotion used to express outrage. Three pre-registered studies ( $N$ 's = 800, 1630, 1100) revealed that people infer different types of aggression from individuals who expressed anger nonverbally compared with those who expressed disgust nonverbally. Perceptions that the outraged individual was angry corresponded with expectations of direct aggression rather than indirect aggression, whereas perceptions that the outraged individual was disgusted corresponded with expectations of indirect aggression rather than direct aggression. These results revealed that the distinct emotions used to communicate outrage shape observers' expectations of how moral conflicts will unfold.

## Keywords

emotional expression, outrage, disgust, anger, aggression

## Introduction

Expressions of moral outrage are ubiquitous, and increasingly so in the online ecosystems that facilitate their rapid transmission. Current thinking suggests that outrage has a wide range of consequences, including (1) motivating people to shame and punish wrongdoers (Salerno & Peter-Hagene, 2013), (2) catalyzing collective action (Spring et al., 2018), and (3) increasing political polarization and intergroup hostility (Brady & Crockett, 2019; Crockett, 2017). Each of these lines of thinking has either treated outrage as a single emotion (e.g., Brady & Crockett, 2019; Crockett, 2017) or interpreted it synonymously with anger (e.g., Spring et al., 2018).

However, outrage comes in different shades, some more resembling anger and others more resembling disgust (e.g., Molho et al., 2017). Comprehensively understanding and harnessing the power of outrage may require a more nuanced understanding of its communicative effects. In particular, is the outrage expressed via anger interchangeable with that expressed via disgust, or do these two methods of communicating outrage have different effects on observers? Here, we propose and test the hypothesis that the specific emotion used to communicate outrage affects how observers infer behavioral responses from the expressor.

## Shades Within Expressions of Outrage

Evolutionary perspectives of emotion argue that anger functions to change the behavior of others — particularly those who are perceived as insufficiently valuing another's welfare (Sell et al., 2017; Szyner et al., 2021, 2022). It does so by motivating direct aggression (e.g., physical violence or verbal threats), which can immediately stop transgressions or disincentivize future ones. Consistent with this perspective, diary and vignette studies indicate that people who experience anger toward moral violations are more likely to feel motivated to directly aggress against moral violators (e.g., Lopez et al., 2021; Molho et al., 2017; Molho et al., 2020; Wyckoff, 2016).

The nature of moral disgust is more mysterious. Some scholars argue that it is merely a synonym for anger (e.g., Alvarado, 1998; Nabi, 2002). Consistent with this view,

<sup>1</sup>Vrije Universiteit Amsterdam, the Netherlands

<sup>2</sup>Nottingham Trent University, UK

<sup>3</sup>University of Amsterdam, the Netherlands

## Corresponding Author:

Lei Fan, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Room MF-B555, Van der Boechorststraat 7, 1081 BT Amsterdam, the Netherlands.

Email: l.fan@vu.nl

verbal self-reports of anger and disgust toward moral violations are highly correlated (e.g.,  $r$ 's = .77 to .82, as reported in Gutierrez and Giner-Sorolla (2007). Even so, some studies using verbal self-reports of anger and disgust report distinct relations between the two emotions and other variables (e.g., Gutierrez et al., 2012; Russell & Giner-Sorolla, 2011). Other studies assessing anger and disgust via endorsements of canonical facial expressions rather than linguistic labels have revealed larger distinctions between the two. For example, findings using this method suggest that disgust, more than anger, is activated in response to information about (bad) moral character (e.g., Giner-Sorolla & Chapman, 2017); that anger communicates more self-interest, whereas disgust communicates a more principled, moral motivation (Kupfer & Giner-Sorolla, 2017); and that disgust is activated more when moral violations target others than when they target the self or relatives (Lopez et al., 2021; Molho et al., 2017; Ocampo et al., 2022). These observations align with proposals that moral disgust functions to coordinate condemnation and, in doing so, share the costs of punishment with other bystanders (e.g., Chapman et al., 2009; Tybur et al., 2013). Consistent with this perspective, moral disgust seems to be associated with more indirectly aggressive responses, such as gossip and ostracism (e.g., Molho et al., 2020).

These considerations raise the following question: Do people expect aggression from the morally outraged and, if so, what kind? We tested the hypothesis that observers infer different degrees of direct and indirect aggression from individuals who express outrage via anger versus disgust. Given some of the ambiguities of verbal expressions of emotion (e.g., Nabi, 2002), we test whether such intentions can be communicated nonverbally.

### ***Non-Verbal Emotional Expressions***

Emotional expressions function to communicate information and incentivize specific social behaviors (Keltner et al., 2022; Shariff & Tracy, 2011; van Kleef, 2016). Although people can convey outrage with mere verbal labels (e.g., “I am currently feeling disgust”), they often do so via facial and vocal expressions — that is, nonverbally. Uncovering distinct effects of the different shades of expressions of outrage might require an approach beyond verbal self-reports of emotion — one that captures the multi-modal communication of emotion via non-verbal vocalizations and facial expressions.

### ***Overview of Studies***

The current investigation aimed to test whether observers anticipate different types of aggressive behaviors from individuals who express outrage toward moral violations via anger versus via disgust. Following recent recommendations in the emotion literature (Weidman et al., 2017), we employed a multi-modal presentation of emotional

expressions that combined non-verbal facial and vocal expressions. In Study 1, participants saw and heard an individual express either disgust or anger toward an unspecified moral violation. In Study 2, the moral violation was described in detail so that the assumed nature of the moral violation was not confounded with the type of nonverbal expressions made by the individual expressing outrage. Study 3 replicated Study 2 while adding a comparison condition in which the individual observing the moral violation produced no nonverbal expression.

Each study used a stimulus sampling approach in which participants read that an expressor had witnessed an unspecified moral violation (Study 1) or read one of 12 moral violations that the expressor witnessed (Studies 2 and 3). They then saw and heard one of 12 stimuli from the expressor, six of which expressed anger and six of which expressed disgust (Studies 1–3). We aimed to test the hypothesis that observers infer different degrees of direct and indirect aggression from individuals who express outrage via anger versus disgust. Specifically, we hypothesized that observers would make stronger inferences of direct aggression relative to indirect aggression from anger expressors relative to disgust expressors — that is, a two-way interaction between type of emotion and type of aggression.

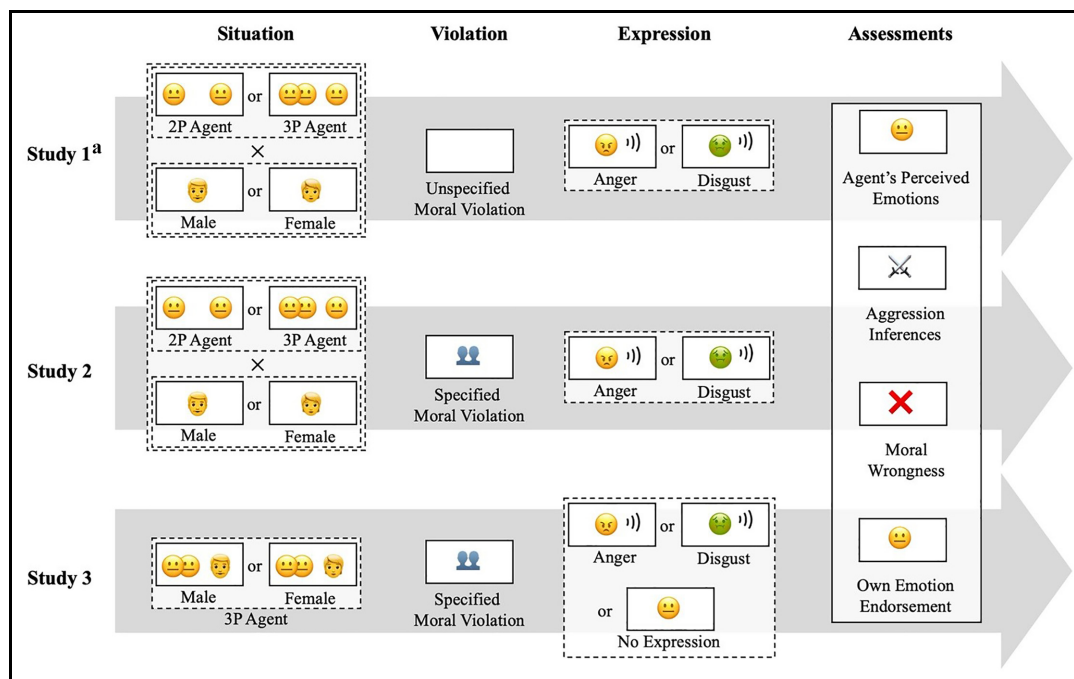
In Studies 1 and 2, we also manipulated whether the expressor was directly affected by or merely witnessed the moral violation, and we also pre-registered hypotheses concerning effects of this manipulation on direct versus indirect aggression. We did not pre-register any hypotheses regarding interactions between second- versus third-party condition and emotional expression. However, we conducted exploratory analyses to evaluate the generalizability of the effect of emotion expression on direct versus indirect aggression across second- and third-party conditions (cf. Kupfer & Giner-Sorolla, 2017). The results showed consistent patterns in both conditions (see supplementary online material (SOM) for details).

Ethical approval of the current project was granted by the Scientific and Ethical Review Board (VCWE) of the Faculty of Behavior & Movement Sciences, VU University Amsterdam (VCWE-2021-032). Across studies, participants were recruited via Prolific. All studies were pre-registered before data collection. Preregistrations, materials, data, and analysis scripts for all three studies are available on OSF (<https://osf.io/6n2qp/>).

## **Study 1**

### ***Method***

***Participants.*** Based on an a priori power analysis with SimR (Green & MacLeod, 2016), we targeted a sample of 700 adult participants to complete two survey sessions, which were separated by approximately 1 week. This sample size afforded 85% power to detect a small to medium interaction effect between aggression inference (direct vs. indirect)



**Figure 1.** Designs and Task Overviews of the Three Studies in the Current Paper

<sup>a</sup>Study 1 used a mixed design, in which each participant completed two sessions of the task.

and emotion expression (see SOM). We recruited 800 native English speakers (34.13% male; age:  $M = 32.15$  years,  $SD = 11.40$ ) from Prolific and invited all valid participants in the first session to continue for the second session. Six hundred forty of these participants completed both sessions. Participants who only provided data in the first session were included in the analyses. Participants provided informed consent before and were paid after their participation.

**Procedures.** In one of the two sessions, an agent was presented as the target of a moral violation (i.e., second-party punishment condition); in the other session, the agent was presented as an observer of a moral violation that targets another individual (i.e., third-party punishment condition). In one of the two sessions, the agent expressed anger, whereas in the other session, the agent expressed disgust. Agent conditions (i.e., second- vs. third-party) and emotion expression conditions were paired randomly for each participant across the two sessions. We asked participants to rate the degree to which the agent expressed each of six emotions and the degree to which they thought the actor would engage in directly aggressive and indirectly aggressive acts toward the moral violator. Participants also reported their sex and age. For an overview of the procedure, see Figure 1.

Participants also rated the moral wrongness of the moral violator's behavior and their own emotional reaction to the moral violation, and they provided free-response descriptions of the last morally wrong act they witnessed (for the

first session) and the last morally disgusting act they witnessed (for the second session). These items were included for exploratory purposes.

### Materials

**Scenario.** Scenarios described an agent (named Olivia or Oliver) as either targeted by an offender (named Mark) or witnessing someone else (named John) targeted by an unspecified moral violation.

**Emotion Expression Stimuli.** We retrieved anger and disgust facial expressions from the subset of front-facing front-gazing White targets from the Radboud Face Database (RaFD, Langner et al., 2010). Based on RaFD validation data, we selected the faces with the highest agreement scores in each emotion of each sex. In all, 12 faces were selected as the final facial expression set, with three faces for each emotion of each sex. Non-verbal vocal tokens used by Sauter et al. (2010) were selected through a pilot study with an emotion recognition task. We selected 12 clips as the final vocal token set, with three tokens for each emotion of each sex.

Stimuli included one facial expression and one vocal expression from the same emotion and the same actor sex. The stimulus pool included three possible stimuli per sex per emotion (a total of six disgust stimuli and six anger stimuli). Participants were presented with one stimulus from the pool according to the emotion expression and agent sex conditions they were assigned to.

**Manipulation Check.** We included a set of manipulation check items in the survey: “To what extent is Olivia/Oliver feeling the following emotions in response to Mark’s behavior?” with items from six emotions (i.e., happiness, sadness, fear, disgust, anger, and surprise) rated on 7-point Likert-type scales (1 = *not at all* to 7 = *extremely*).

**Aggression Inferences.** We adapted items assessing aggression motivations from Molho et al. (2017). We included five direct aggression items (e.g., “Olivia/Oliver would hit Mark”) and six indirect aggression items (e.g., “Olivia/Oliver would spread negative information about Mark”) asking how well each statement describes the agent’s likely behavior on a 7-point Likert-type scale (1 = *extremely unlikely*, 7 = *extremely likely*).

**Analytic Approach.** Using mixed-effects modeling, we regressed dependent measures on emotion expression condition (anger vs. disgust), agent condition (second- vs. third-party), agent sex (male vs. female), participant sex (male vs. female), and participant age.

As pre-registered, we tested effects of emotion expression conditions on inferences of aggression and perceived emotions through interactions in the models. For models focusing on inferences of aggression, we added the type of aggression (direct vs. indirect) and targeted interactions (i.e., type of aggression by emotion expression condition; type of aggression by agent condition) as predictors. For models focusing on perceived emotions, we added the type of emotion (anger vs. disgust) and targeted interactions (i.e., type of perceived emotions by emotion expression condition; type of perceived emotions by agent condition) as predictors.

We also report analyses we did not pre-register: the effect of perceptions of anger and disgust on inferences of the two aggression types. In these analyses, we added the two emotion perceptions and targeted interactions (i.e., type of aggression by emotion perceptions) into the model predicting aggression inferences. Tests of additional exploratory hypotheses are described in the SOM.

We followed modeled random intercepts for stimuli nested within agent sex and emotion condition, random intercepts for participants, and random slopes for emotion condition and agent condition within participants. We then conducted simple-effect analyses within the interactions.

## Results

**Data Exclusion.** We pre-registered excluding respondents with nonsensical answers to two free-response questions. No participants were excluded based on these criteria.

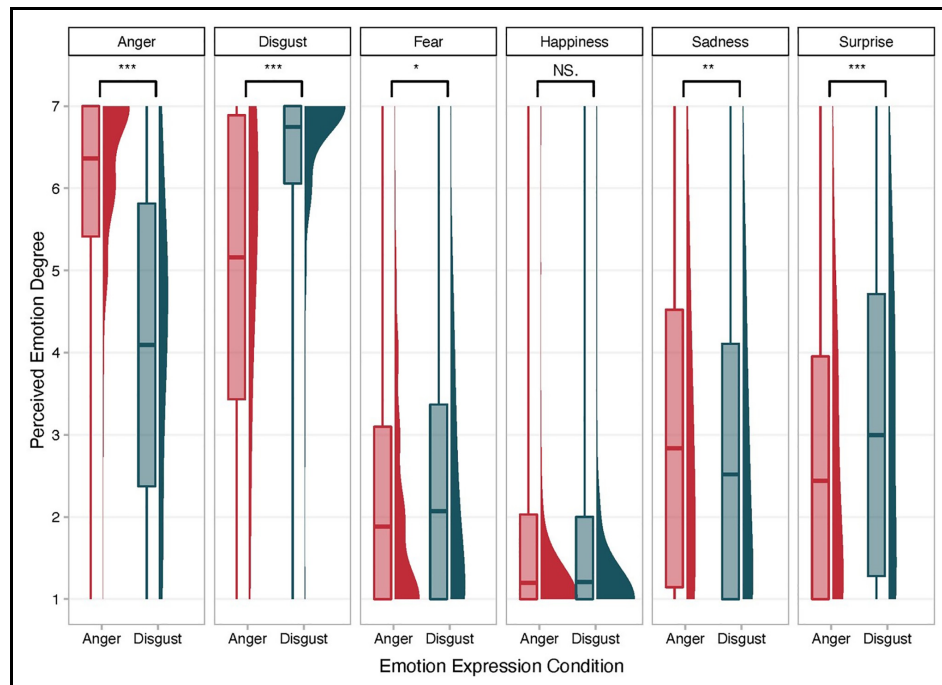
**Manipulation Check.** Across conditions, participants most strongly perceived anger and disgust, and they did so in a way consistent with the manipulations: Perceptions of

anger were higher in the anger expression condition ( $M = 6.36$ , 95% confidence interval [CI] [6.23, 6.50]) than in the disgust expression condition ( $M = 4.09$ , 95% CI [3.96, 4.22]),  $t(25.4) = 26.37$ ,  $p < .001$ , and perceptions of disgust were higher in the disgust expression condition ( $M = 6.74$ , 95% CI [6.61, 6.88]) than in the anger expression condition ( $M = 5.16$ , 95% CI [5.03, 5.29]),  $t(25.4) = 18.36$ ,  $p < .001$ . In each expression condition, the corresponding target emotion was perceived stronger than the other emotion,  $t_{\text{Disgust}}(7824) = 41.49$ ,  $p < .001$ ;  $t_{\text{Anger}}(7824) = 18.71$ ,  $p < .001$ . Perceptions of the other emotions were much lower, and they varied more modestly across conditions (see Figure 2).

**Effect of Emotion Expression Manipulation.** Consistent with hypotheses, the interaction between emotion expression condition and aggression type was statistically significant,  $\beta = .19$ ,  $F(1, 1437) = 217.73$ ,  $p < .001$ , 95% CI [.16, .21] (see Figure 3A). Participants inferred more direct aggression in the anger condition ( $M = 4.57$ , 95% CI [4.37, 4.78]) than in the disgust condition ( $M = 3.63$ , 95% CI [3.42, 3.83]),  $t(9.48) = 7.42$ ,  $p < .001$ , but there was no difference in inferences of indirect aggression across the two conditions ( $M_{\text{Anger}} = 4.18$ , 95% CI [3.97, 4.38],  $M_{\text{Disgust}} = 4.22$ , 95% CI [4.01, 4.42]),  $t(9.48) = -0.31$ ,  $p = .77$ . In the anger expression condition, direct aggression inferences were higher than indirect aggression inferences,  $t(1437) = 8.36$ ,  $p < .001$ ; in the disgust expression condition, direct aggression inferences were lower than indirect aggression inferences,  $t(1437) = -12.52$ ,  $p < .001$ .

**Effect of Agent Manipulation.** We followed the same pipeline to test whether expressor’s agent roles affect inferences of direct aggression versus indirect aggression. The interaction was not statistically significant,  $\beta = .02$ ,  $F(1, 1437) = 2.84$ ,  $p = .09$ , 95% CI [-.00, .05].

**Effect of Perceived Emotions.** The interaction between aggression type and participants’ perceptions of the agent’s emotions was significant for both anger ( $\beta = .17$ ,  $p < .001$ , 95% CI [.14, .20]) and disgust ( $\beta = -.07$ ,  $p < .001$ , 95% CI [-.09, -.04]). Perceptions of anger more strongly related to inferences of direct aggression, ( $\beta = .36$ , 95% CI [.32, .40]) than indirect aggression ( $\beta = .11$ , 95% CI [.07, .15]),  $t(1437) = 13.06$ ,  $p < .001$ . Conversely, perceptions of disgust related more strongly to inferences of indirect aggression ( $\beta = .10$ , 95% CI [.05, .13]) than to inferences of direct aggression ( $\beta = -.02$ , 95% CI [-.07, .02]),  $t(1435) = 7.41$ ,  $p < .001$ . All effects remained when controlling for the main effects of the two manipulations and their interaction with aggression type — that is, even after controlling for the type of stimulus (anger or disgust) participants saw.



**Figure 2.** Distributions of Perceived Emotions Between Emotion Expression Conditions in Study 1

Note. In each boxplot, the horizontal line indicates the mean, the box indicates the range between one standard deviation above and below the mean, and the vertical line indicates the full range of responses. The shaded area indicates the density of the data after smoothing. Significance signs indicate the inter-conditional simple-effect comparisons.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

All the other figures in this article follow the same indication methods.

## Study 2

In Study 1, we did not describe the moral violation that the agent witnessed or was targeted by. This approach might have led participants to assume that anger expressors witnessed different moral violations than disgust expressors did. For example, participants might have been more likely to assume that disgust expressors witnessed moral violations with some sexual or pathogen content (i.e., purity violations), whereas anger-expressors witnessed moral violations with some harm or unfairness content (e.g., Heerdink et al., 2019; Kupfer et al., 2020; Rozin et al., 1999). In Study 2, we eliminated this alternative explanation by describing the content of the moral violation that agents responded to.

### Method

**Participants.** Based on an a priori power analysis, we targeted a sample of 1,600 participants, which afforded 85% power to detect the interaction between inferred aggression type (direct vs. indirect) and emotional expression observed in Study 1 (for full details, see SOM). We recruited 1,630 native English speakers (50.06% male; age:  $M = 41.01$  years,  $SD = 14.55$ ) residing in the United Kingdom from Prolific. Participants provided informed consent before and were paid after their participation.

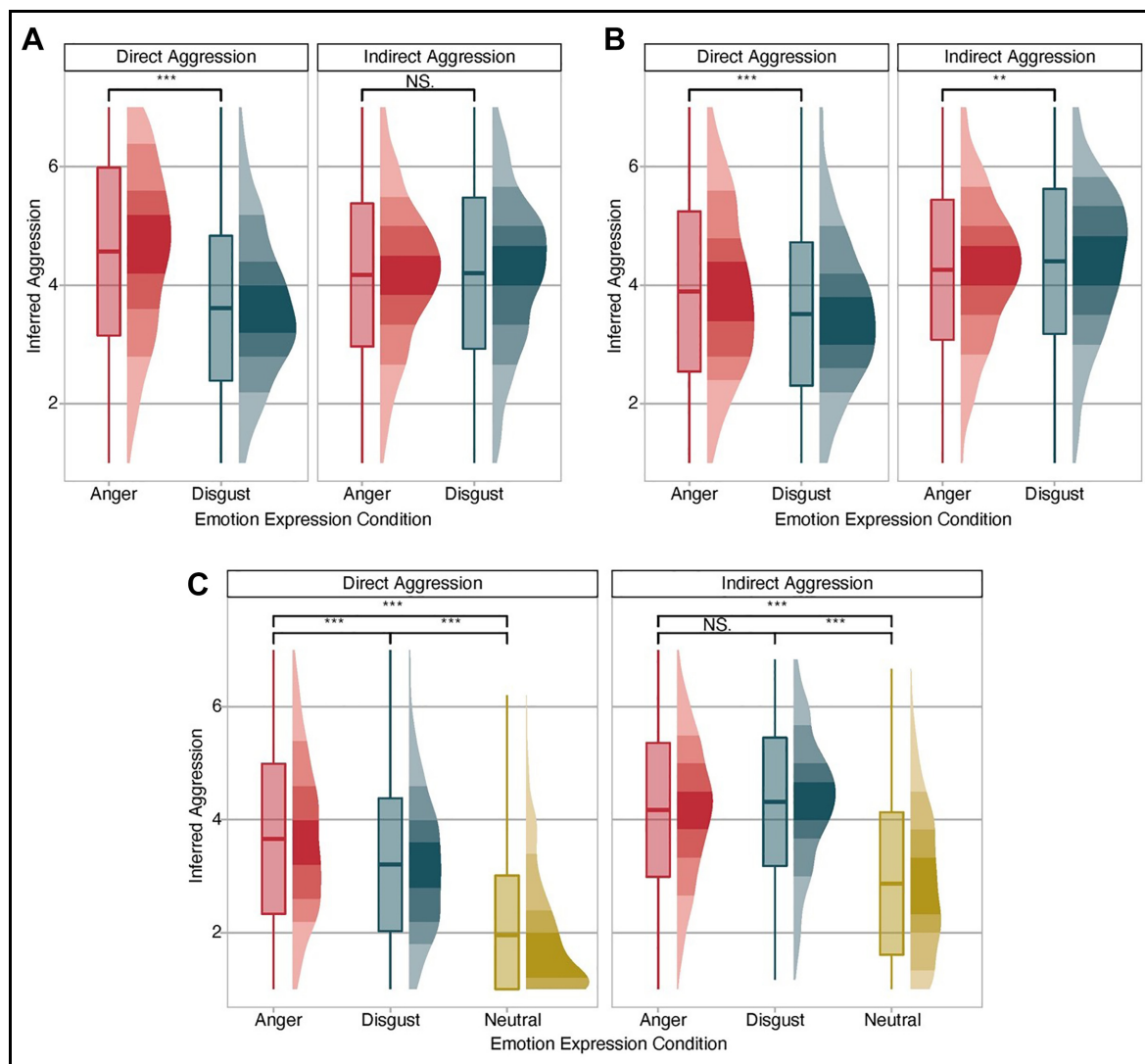
**Procedures.** Given that approximately 20% of Study 1 participants did not complete both sessions, we employed a single-session between-subject design in Study 2. Participants were randomly assigned to one of the four conditions used in Study 1 (i.e., second- vs. third-party expressors by disgust vs. anger expressions).

Participants first read a scenario describing an agent who was either directly affected by or witnessed a moral violation and responded nonverbally by expressing either anger or disgust. They then rated the degree to which the agent expressed the same six emotions as in Study 1 and the degree to which they perceived the agent as likely to engage in directly and indirectly aggressive acts against the moral violator. Participants also rated the moral wrongness of the moral violator's behavior and their own emotional reactions to the moral violator's behavior. Finally, they reported their own sex and age.

### Materials

**Moral Violations.** We generated 12 scenarios inspired by the moral violation scenarios that the participants in Study 1 reported recently witnessing in their own lives (further details see SOM). The scenarios were adapted to the agent role (second- or third-party) and agent sex.

**Analytic Approach.** The analytic approach was identical to that used in Study 1, with one exception: We modeled



**Figure 3.** Distributions of Aggression Inference of the Two Aggression Types Between Emotional Expression Conditions in Study 1(C), Study 2(B), and Study 3(C)

random intercepts for moral violation scenarios and random slopes of emotion type and agent role (second- or third-party) across scenarios.

## Results

**Data Exclusion.** We pre-registered excluding respondents with nonsensical answers to the free-response question: “What do you think about the emotional response of Oliver/Olivia in the scenario?” No participants were excluded based on this criterion.

**Manipulation Check.** We used the same pipeline for the manipulation checks from Study 1. The results were consistent with those in Study 1 (see SOM).

**Effect of Emotion Expression Manipulations.** Consistent with Study 1, the interaction between aggression type and

emotion expression condition was statistically significant,  $\beta = .10$ ,  $F(1, 1626) = 81.30$ ,  $p < .001$ , 95% CI [.08, .12] (see Figure 3B). Participants inferred more direct aggression when the agent expressed anger ( $M = 3.90$ , 95% CI [3.73, 4.07]) than when the agent expressed disgust ( $M = 3.51$ , 95% CI [3.33, 3.68]),  $t(10.5) = 3.67$ ,  $p = .004$ , but no difference was observed for indirect aggression across the two conditions, ( $M_{\text{Anger}} = 4.27$ , 95% CI [4.09, 4.44],  $M_{\text{Disgust}} = 4.39$ , 95% CI [4.22, 4.57],  $t(10.5) = -1.21$ ,  $p = .25$ ). In both the anger and disgust expression conditions, participants inferred less direct aggression than indirect aggression,  $t_{\text{Anger}}(1626) = -8.97$ ,  $p < .001$ ,  $t_{\text{Disgust}}(1626) = -21.68$ ,  $p < .001$ , though the difference was larger when disgust was expressed than when anger was expressed.

**Effect of Agent Manipulations.** The interaction between agent condition and aggression type was significant,  $\beta = .03$ ,

$F(1, 1626) = 9.20, p = .003, 95\% \text{ CI } [.01, .16]$ . Participants inferred more direct aggression from second parties ( $M = 3.82, 95\% \text{ CI } [3.69, 3.96]$ ) than from third parties,  $M = 3.58, 95\% \text{ CI } [3.45, 3.72], t(2519) = 4.08, p < .001$ , but no differences for indirect aggression across the two conditions,  $M_{2P} = 4.36, 95\% \text{ CI } [4.23, 4.50], M_{3P} = 4.30, 95\% \text{ CI } [4.16, 4.43], t(2519) = 1.13, p = .26$ .

**Effect of Perceived Emotions.** The interaction between aggression type and perception of agent's emotion was significant for both anger ( $\beta = .08, p < .001, 95\% \text{ CI } [.06, .10]$ ) and disgust ( $\beta = -.03, p = .004, 95\% \text{ CI } [-.05, -.01]$ ). Perceptions of anger more strongly related to inferences of direct aggression ( $\beta = .29, 95\% \text{ CI } [.24, .33]$ ) than indirect aggression ( $\beta = .14, 95\% \text{ CI } [.10, .19]$ ),  $t(1626) = 7.15, p < .001$ , whereas perceptions of disgust more strongly related to inferences of indirect aggression ( $\beta = .13, 95\% \text{ CI } [.09, .17]$ ) than direct aggression ( $\beta = .07, 95\% \text{ CI } [.02, .11]$ ),  $t(1626) = -2.86, p = .004$ .

### Study 3

The first two studies suggested that perceivers expect different types of aggression from the morally outraged depending on their emotion expression. However, observers might infer similar aggression — perhaps especially indirect aggression — from individuals who merely witness a moral violation but do not express emotion. Such a possibility would undercut the information value of emotional expressions of outrage. Thus, we aimed for a further test of the effect of disgust expressions on expectations of indirect aggression beyond inferences when no emotion expressed is expressed.

#### Method

**Participants.** Based on an a priori power analysis using the effect sizes observed in Study 2 and lower expectations of aggression from non-expressors, we targeted a sample of 1100 participants, which afforded 85% power to detect the interaction between emotional expression and inferred aggression type (direct vs. indirect; for full details see SOM). Given the low exclusion rate in Studies 1 and 2, we recruited 1100 native English speakers (50.05% male; age:  $M = 39.42$  years,  $SD = 13.46$ ), residing in the United Kingdom, from Prolific. Participants were paid £0.75. All respondents provided informed consent.

**Procedures.** Procedures were identical to those used in Study 2, except that all agents were third-party observers of a moral violation.

#### Materials

**Emotion Expression.** For the anger and disgust conditions, we used the same emotion expression stimuli as in Studies

1 and 2. For the neutral expression condition, a face expressing no emotion was paired with text stating that the agent's facial expression did not change in response to the violator's actions, and the agent did not make any sound. The neutral facial expressions were also retrieved from the RaFD (Langner et al., 2010).

**Manipulation Check.** We included the same items used in Studies 1 and 2 and added two 7-point items assessing response valence ("How negative versus positive was Olivia's/Oliver's reaction to Mark's behavior?") and intensity ("How intense was Olivia's/Oliver's reaction to Mark's behavior?") (for valence,  $-3 =$  extremely negative,  $3 =$  extremely positive; for intensity,  $0 =$  not intense at all,  $6 =$  extremely intense). These items were intended as a manipulation check for the non-expression condition.

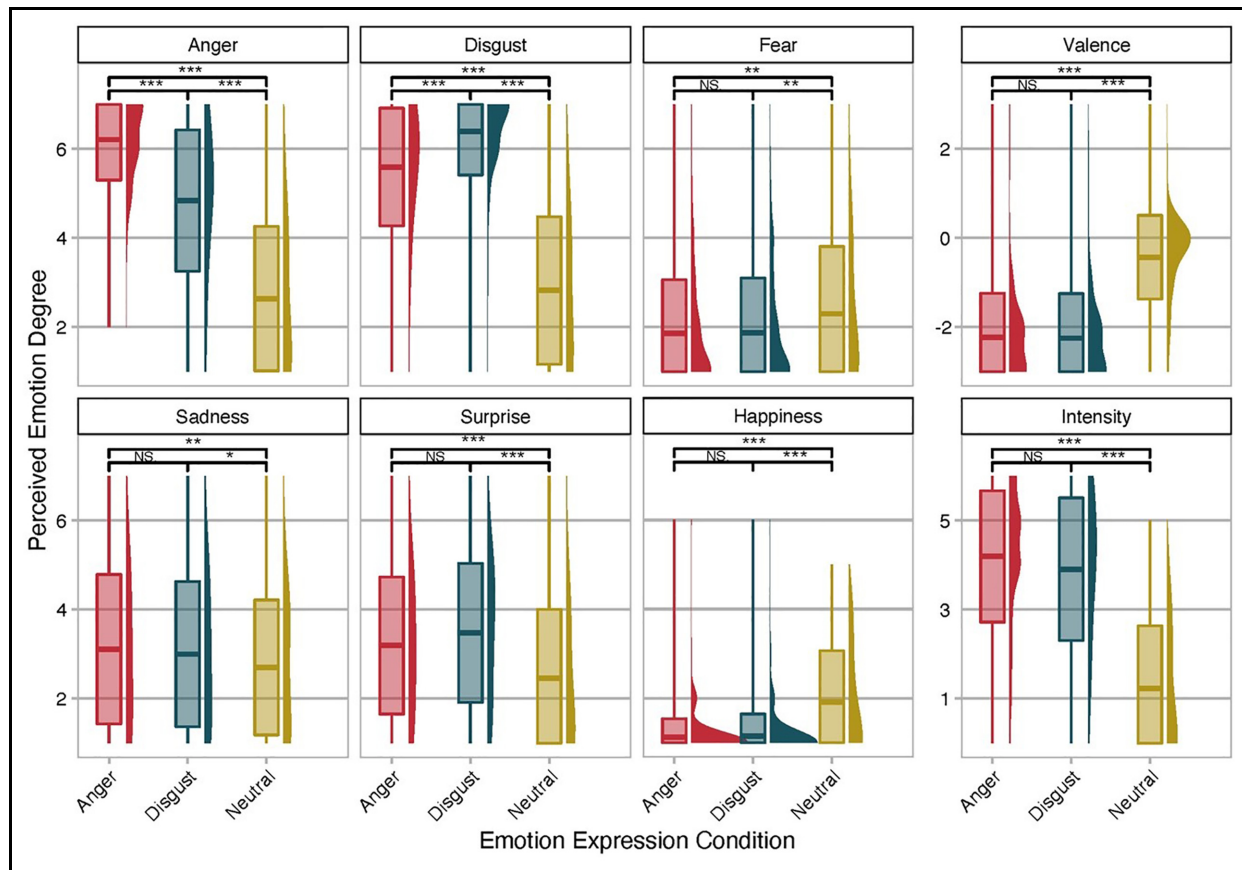
**Analytic Approach.** We followed the same analytic approach used in Study 2.

#### Results

**Data Exclusion.** We pre-registered excluding respondents with nonsensical answers to the free-response question. No participants were excluded based on this criterion.

**Manipulation Check.** As in Studies 1 and 2, participants perceived anger and disgust in a manner consistent with the manipulations (see Figure 4, details see SOM). We detected no difference of perceptions of response valence across anger ( $M = -2.24, 95\% \text{ CI } [-2.35, -2.13]$ ) and disgust conditions ( $M = -2.25, 95\% \text{ CI } [-2.36, -2.14]$ ),  $t(1083) = 0.20, p = .98$  ( $p$  value was adjusted using Tukey's method for comparing the variable of expression condition with 3 levels of manipulations, same below), though expressions in both conditions were perceived as more negative than those in the neutral condition,  $M = -0.43, 95\% \text{ CI } [-0.54, -0.32], t_{\text{Anger}}(1083) = -25.15, p < .001, t_{\text{Disgust}}(1083) = -25.35, p < .001$ . We observed the same pattern for perceptions of response intensity; no difference was detected between anger ( $M = 4.20, 95\% \text{ CI } [3.89, 4.51]$ ) and disgust ( $M = 3.91, 95\% \text{ CI } [3.60, 4.22]$ ),  $t(14) = 1.42, p = .36$  conditions, but expressions in both conditions were perceived as more intense compared with those in the neutral condition,  $M = 1.22, 95\% \text{ CI } [0.91, 1.53], t_{\text{Anger}}(14) = 14.54, p < .001, t_{\text{Disgust}}(14) = 13.13, p < .001$ .

**Effect of Emotion Expression Manipulations.** Consistent with Studies 1 and 2, we observed an interaction between emotion expression condition and aggression type on expectations of aggression,  $F(2, 1096) = 26.80, p < .001$  (see Figure 3C). Again, participants inferred more direct aggression when the agent expressed anger ( $M = 3.65, 95\% \text{ CI } [3.47, 3.82]$ ) than when the agent expressed disgust ( $M =$



**Figure 4.** Distributions of Perceived Emotions, Response Valence, and Intensity Between Emotion Expression Conditions in Study 3

3.21, 95% CI [3.04, 3.39]),  $t(18.7) = 3.77, p = .004$ , but there was no difference in inferences of indirect aggression across the two conditions,  $M_{\text{Anger}} = 4.16$ , 95% CI [3.98, 4.33],  $M_{\text{Disgust}} = 4.33$ , 95% CI [4.15, 4.50],  $t(18.7) = -1.47, p = .33$ . When the agent did not express emotion, participants inferred less direct ( $M = 1.96$ , 95% CI [1.78, 2.13]) and indirect aggression ( $M = 2.87$ , 95% CI [2.70, 3.05]) than when the agent expressed either anger,  $t_{\text{DA}}(18.6) = 14.81, p < .001$ ,  $t_{\text{IA}}(18.6) = -11.26, p < .001$  or disgust,  $t_{\text{DA}}(18.6) = 11.04, p < .001$ ,  $t_{\text{IA}}(18.6) = 12.74, p < .001$ .

**Effect of Perceived Emotions.** The interaction between aggression type and perception of agent's emotion was significant for both anger ( $\beta = .12, p < .001$ , 95% CI [.08, .15]) and disgust ( $\beta = -.10, p < .001$ , 95% CI [-.13, -.07]). Perceptions of anger more strongly related to inferences of direct aggression ( $\beta = .32$ , 95% CI [.26, .37]) than indirect aggression ( $\beta = .15$ , 95% CI [.10, .20]),  $t(1096) = 7.05, p < .001$ , whereas perceptions of disgust more strongly related to inferences of indirect aggression ( $\beta = .17$ , 95% CI [.11, .22]) than direct aggression ( $\beta = .02$ , 95% CI [-.03, .08]),  $t(1096) = -6.13, p < .001$ . All results remained statistically significant when controlling for the main

effects of the two manipulations and their interactions with aggression type.

## Discussion

Do people expect aggression from the morally outraged? Results consistently suggest that they do, but that the nature of the anticipated aggression depends on the emotion used to express outrage. Across three studies, observers expected greater direct aggression from individuals who nonverbally expressed anger relative to those who nonverbally expressed disgust, but no differences in indirect aggression across anger and disgust expressors. However, exploratory tests revealed the following pattern in all three studies: The degree to which disgust was perceived from an outraged individual related more strongly to expectations of indirect aggression than direct aggression, whereas the degree of perceived anger related more strongly to expectations of direct aggression than indirect aggression. These results have implications for how we understand moral emotions and the consequences of outrage.

### Implications for Understanding Moral Emotions

The nature — and even existence — of moral disgust has been debated for decades (Chapman & Anderson, 2013;



Nabi, 2002; Royzman & Kurzban, 2011; Rozin et al., 1999). Such debates have largely focused on whether the qualia (i.e., subjective experience) and action tendencies associated with disgust toward moral violations are identical to those associated with disgust toward pathogen cues (e.g., bodily waste, spoiled foods). Less work has focused on understanding the distinctions between angry and disgusted reactions to moral violations. The current findings indicate that people infer different action tendencies from expressors of disgust versus anger, and they do so in a way that is consistent with self-reports of anger versus disgust on motivations to directly versus indirectly aggress (Hutcherson & Gross, 2011; Molho et al., 2017; Molho et al., 2020; Ocampo et al., 2022). All together, these results raise multiple intriguing possibilities, including that individuals might calibrate the degree of anger versus disgust in their expressions to strategically shape others' expectations of their behaviors. Such calibration could afford multiple benefits. For example, given the reputational costs of direct aggression (e.g., Eriksson et al., 2016), individuals could express disgust rather than anger when such costs are high (see also Kupfer & Giner-Sorolla, 2017). In a similar vein, low-power individuals, who have been found to moralize more (Petersen, 2013; Wang & Inbar, 2021), might more frequently express such moralization via disgust so as to not provoke counter aggression that anger might elicit. As another example, if disgust expressions communicate a lower probability of direct intervention, then they may be used to reduce the likelihood of second-order free-riding (i.e., taking advantage of others' punishment of the transgressor) and increase the likelihood of coordinated punishment across individuals.

These considerations can also extend current thinking on the effects of outrage expressed in dense social networks — perhaps particularly online ones. Much like research on moral emotions in general, research into (online) outrage has typically collapsed across distinct emotions. Outrage expressions might boost collective responses toward social norm violations via the potential condemnation-coordinating function of disgust. However, they might also increase polarization, conspiracy beliefs, and radicalization through the stronger tendency toward violence that is characteristic of anger (cf. Brady & Crockett, 2019; Crockett, 2017; Spring et al., 2018). The current results indicate that anger and disgust expressions might lead to different behavioral inferences, which may lead to different consequences. When observing anger expressions or perceiving more anger, people would infer a high likelihood of direct aggressive actions by the expressor. This might in turn shape response strategies by the targets of the potential punishment. The stronger inference of direct aggression can be regarded as a violent threat that may also boost retaliation and lead to aggression escalation. Added to the low cost of online expression in the current era, anger expressions might be more dangerous than disgust expressions in affecting our social interactions.

## Limitations and Directions for Future Research

The findings are limited by a few factors, each of which provides opportunities for future research. First, in each study, outrage was expressed toward an adult White male target. Some evidence suggests that individuals are more willing to punish male moral violators than female ones (e.g., Reynolds et al., 2020). Further work is needed to test whether the current findings generalize across situations in which the moral violator has other demographic characteristics. Second, all studies enrolled only English-speaking participants, and Studies 2 and 3 only enrolled participants from the United Kingdom. Vocal communication of socially disengaging emotions (perhaps especially anger) varies across cultures, possibly as a result of differences in the cultural norms of interpersonal relationships between societies (Yoshie & Sauter, 2020). Future research could compare the effects of emotional expressions from different cultures and test the extent to which the present results replicate across different societies. In the meantime, the field will benefit from generating and testing other hypotheses that can illuminate the functions of moral emotions, particularly in terms of punishment.

## Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project is funded by China Scholarship Council (201806990045).

## ORCID iDs

Lei Fan  <https://orcid.org/0000-0002-3531-9199>

Catherine Molho  <https://orcid.org/0000-0002-2846-2544>

Joshua M. Tybur  <https://orcid.org/0000-0002-0462-6508>

## Supplemental Material

Supplemental material for this article is available at <https://doi.org/10.1177/19485506231176954>

## References

- Alvarado, N. (1998). A reconsideration of the structure of the emotion lexicon. *Motivation and Emotion, 22*(4), 329–344. <https://doi.org/10.1023/A:1021356424065>
- Brady, W. J., & Crockett, M. J. (2019). How effective is online outrage? *Trends in Cognitive Sciences, 23*(2), 79–80. <https://doi.org/10.1016/j.tics.2018.11.004>
- Chapman, H. A., & Anderson, A. K. (2013). Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological Bulletin, 139*(2), 300–327. <https://doi.org/10.1037/a0030964>

- Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222–1226. <https://doi.org/10.1126/science.1165565>
- Crockett, M. J. (2017). Moral outrage in the digital age. *Nature Human Behaviour*, *1*(11), 769–771. <https://doi.org/10.1038/s41562-017-0213-3>
- Eriksson, K., Andersson, P. A., & Strimling, P. (2016). Moderators of the disapproval of peer punishment. *Group Processes & Intergroup Relations*, *19*(2), 152–168. <https://doi.org/10.1177/1368430215583519>
- Giner-Sorolla, R., & Chapman, H. A. (2017). Beyond purity: Moral disgust toward bad character. *Psychological Science*, *28*(12), 1871–1871. <https://doi.org/10.1177/0956797616673193>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Gutierrez, R., & Giner-Sorolla, R. (2007). Anger, disgust, and presumption of harm as reactions to taboo-breaking behaviors. *Emotion*, *7*(4), 853. <https://doi.org/10.1037/1528-3542.7.4.853>
- Gutierrez, R., Giner-Sorolla, R., & Vasiljevic, M. (2012). Just an anger synonym? Moral context influences predictors of disgust word use. *Cognition & Emotion*, *26*(1), 53–64. <https://doi.org/10.1080/02699931.2011.567773>
- Heerdink, M. W., Koning, L. F., van Doorn, E. A., & van Kleef, G. A. (2019). Emotions as guardians of group norms: Expressions of anger and disgust drive inferences about autonomy and purity violations. *Cognition & Emotion*, *33*(3), 563–578. <https://doi.org/10.1080/02699931.2018.1476324>
- Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social-functional account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*(4), 719–737. <https://doi.org/10.1037/a0022408>
- Keltner, D., Sauter, D., Tracy, J. L., Wetchler, E., & Cowen, A. S. (2022). How emotions, relationships, and culture constitute each other: Advances in social functionalist theory. *Cognition and Emotion*, *36*(3), 388–401. <https://doi.org/10.1080/02699931.2022.2047009>
- Kupfer, T. R., & Giner-Sorolla, R. (2017). Communicating moral motives: The social signaling function of disgust. *Social Psychological and Personality Science*, *8*(6), 632–640. <https://doi.org/10.1177/1948550616679236>
- Kupfer, T. R., Inbar, Y., & Tybur, J. M. (2020). Reexamining the role of intent in moral judgements of purity violations. *Journal of Experimental Social Psychology*, *91*, 104043. <https://doi.org/10.1016/j.jesp.2020.104043>
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388. <https://doi.org/10.1080/02699930903485076>
- Lopez, L. D., Moorman, K., Schneider, S., Baker, M. N., & Holbrook, C. (2021). Morality is relative: Anger, disgust, and aggression as contingent responses to sibling versus acquaintance harm. *Emotion*, *21*(2), 376–390. <https://doi.org/10.1037/emo0000707>
- Molho, C., Tybur, J. M., Güler, E., Balliet, D., & Hofmann, W. (2017). Disgust and anger relate to different aggressive responses to moral violations. *Psychological Science*, *28*(5), 609–619. <https://doi.org/10.1177/0956797617692000>
- Molho, C., Tybur, J. M., Van Lange, P. A., & Balliet, D. (2020). Direct and indirect punishment of norm violations in daily life. *Nature Communications*, *11*(1), 1–9. <https://doi.org/10.1038/s41467-020-17286-2>
- Nabi, R. L. (2002). The theoretical versus the lay meaning of disgust: Implications for emotion research. *Cognition & Emotion*, *16*(5), 695–703. <https://doi.org/10.1080/02699930143000437>
- Ocampo, D., Sullivan, J., Dayer, A., Palka, E., Betschart, N., & Holbrook, C. (2022). Prosocial aggression tracks genetic relatedness distinctly from emotional closeness. *Emotion*. Advance online publication. <https://doi.org/10.1037/emo0001175>
- Petersen, M. B. (2013). Moralization as protection against exploitation: Do individuals without allies moralize more? *Evolution and Human Behavior*, *34*(2), 78–85. <https://doi.org/10.1016/j.evolhumbehav.2012.09.006>
- Reynolds, T., Howard, C., Sjastad, H., Zhu, L. K., Okimoto, T. G., Baumeister, R. F., Aquino, K., & Kim, J. (2020). Man up and take it: Gender bias in moral typecasting. *Organizational Behavior and Human Decision Processes*, *161*, 120–141.
- Royzman, E., & Kurzban, R. (2011). Facial movements are not goosebumps: A response to Chapman and Anderson. *Emotion Review*, *3*(3), 274–275. <https://doi.org/10.1177/1754073911402402>
- 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*(4), 574–586. <https://doi.org/10.1037/0022-3514.76.4.574>
- Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, *11*(2), 233–240. <https://doi.org/10.1037/a0022598>
- Salerno, J. M., & Peter-Hagene, L. C. (2013). The interactive effect of anger and disgust on moral outrage and judgments. *Psychological Science*, *24*(10), 2069–2078. <https://doi.org/10.1177/0956797613486988>
- Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual cues in nonverbal vocal expressions of emotion. *Quarterly Journal of Experimental Psychology*, *63*(11), 2251–2272. <https://doi.org/10.1080/17470211003721642>
- Sell, A., Sznycer, D., Al-Shawaf, L., Lim, J., Krauss, A., Feldman, A., Rascanu, R., Sugiyama, L., Cosmides, L., & Tooby, J. (2017). The grammar of anger: Mapping the computational architecture of a recalibrational emotion. *Cognition*, *168*, 110–128. <https://doi.org/10.1016/j.cognition.2017.06.002>
- Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395–399. <https://doi.org/10.1177/0963721411424739>
- Spring, V. L., Cameron, C. D., & Cikara, M. (2018). The upside of outrage. *Trends in Cognitive Sciences*, *22*(12), 1067–1069. <https://doi.org/10.1016/j.tics.2018.09.006>
- Sznycer, D., Sell, A., & Dumont, A. (2022). How anger works. *Evolution and Human Behavior*, *43*(2), 122–132. <https://doi.org/10.1016/j.evolhumbehav.2021.11.007>
- Sznycer, D., Sell, A., & Lieberman, D. (2021). Forms and functions of the social emotions. *Current Directions in Psychological Science*, *30*(4), 292–299. <https://doi.org/10.1177/09637214211007451>
- Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and structure. *Psychological Review*, *120*(1), 65–84. <https://doi.org/10.1037/a0030778>

- van Kleef, G. A. (2016). *The interpersonal dynamics of emotion*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107261396>
- Wang, S. Y. N., & Inbar, Y. (2021). Moral-language use by US political elites. *Psychological Science*, 32(1), 14–26. <https://doi.org/10.1177/0956797620960397>
- Weidman, A. C., Steckler, C. M., & Tracy, J. L. (2017). The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. *Emotion*, 17(2), 267–295. <https://doi.org/10.1037/emo0000226>
- Wyckoff, J. P. (2016). Aggression and emotion: Anger, not general negative affect, predicts desire to aggress. *Personality and Individual Differences*, 101, 220–226. <https://doi.org/10.1016/j.paid.2016.06.001>
- Yoshie, M., & Sauter, D. A. (2020). Cultural norms influence nonverbal emotion communication: Japanese vocalizations of socially disengaging emotions. *Emotion*, 20, 513–517. <https://doi.org/10.1037/emo0000580>

### Author Biographies

**Lei Fan**, presently a PhD candidate and Research Associate at Vrije Universiteit Amsterdam, directs his research on moral emotions, punishments, human behavioral immune system, and intergroup biases.

**Catherine Molho**, an Assistant Professor at Vrije Universiteit Amsterdam, obtained her PhD in social psychology from the same institute. Her research focuses on cooperation, trust, emotions, and punishment.

**Tom R. Kupfer**, a Senior Lecturer at Nottingham Trent University, obtained his PhD in social psychology from the University of Kent. His research encompasses emotions, moral psychology, and gender relations.

**Disa A. Sauter**, an Associate Professor at the University of Amsterdam, obtained her PhD in psychology and cognitive science at University College London. She specializes in studying the communication of emotions via non-verbal expressions.

**Joshua M. Tybur**, a Professor at Vrije Universiteit Amsterdam, obtained his PhD in psychology from the University of New Mexico. His research interests include emotions, moral psychology, and the psychology of infectious disease.

Handling Editor: Gervais Will