



## UvA-DARE (Digital Academic Repository)

### Emotions as guardians of group norms

*Expressions of anger and disgust drive inferences about autonomy and purity violations*

Heerdink, M.W.; Koning, L.F.; van Doorn, E.A.; van Kleef, G.A.

#### DOI

[10.1080/02699931.2018.1476324](https://doi.org/10.1080/02699931.2018.1476324)

#### Publication date

2019

#### Document Version

Final published version

#### Published in

Cognition & Emotion

#### License

CC BY-NC-ND

[Link to publication](#)

#### Citation for published version (APA):

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>

#### General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Emotions as guardians of group norms: expressions of anger and disgust drive inferences about autonomy and purity violations

Marc W. Heerdink , Lukas F. Koning, Evert A. van Doorn and Gerben A. van Kleef 

Department of Social Psychology, University of Amsterdam, Amsterdam, the Netherlands

## ABSTRACT

Other people's emotional reactions to a third person's behaviour are potentially informative about what is appropriate within a given situation. We investigated whether and how observers' inferences of such injunctive norms are shaped by expressions of anger and disgust. Building on the moral emotions literature, we hypothesised that angry and disgusted expressions produce relative differences in the strength of autonomy-based versus purity-based norm inferences. We report three studies (plus three supplementary studies) using different types of stimuli (vignette-based, video clips) to investigate how emotional reactions shape norms about potential norm violations (eating snacks, drinking alcohol), and contexts (groups of friends, a university, a company). Consistent with our theoretical argument, the results indicate that observers use others' emotional reactions not only to infer *whether* a particular behaviour is inappropriate, but also *why* it is inappropriate: because it primarily violates autonomy standards (as suggested relatively more strongly by expressions of anger) or purity standards (as suggested relatively more strongly by expressions of disgust). We conclude that the social functionality of emotions in groups extends to shaping norms based on moral standards.

## ARTICLE HISTORY

Received 11 September 2017  
Revised 3 May 2018  
Accepted 9 May 2018

## KEYWORDS

Social norms; emotional influence; moral emotions; anger; disgust

When individuals enter social groups, they enter a period of socialisation during which they learn to fit in with the group's specific social structure (Levine & Moreland, 1994). An important part of this socialisation process consists of learning the social norms of the group, which are the shared expectations about how group members ought to behave (Cialdini, Reno, & Kallgren, 1990). An accurate understanding of these norms is imperative because conforming to them increases the likelihood of being accepted as a full member (Levine & Moreland, 1994). Prior research has shown that social norms are learnt from many sources of information, including media exposure (Paluck, 2009), visible traces of other's behaviour (Cialdini et al., 1990), and the behaviour of relevant others (Paluck & Shepherd, 2012). In this article, we focus on others' emotional reactions to behaviour as a source of

normative information. Specifically, we draw on Emotions as Social Information (EASI) theory (Van Kleef, 2016; Van Kleef, Van Doorn, Heerdink, & Koning, 2011) and the literature on moral emotions (e.g. Horberg, Oveis, & Keltner, 2011; Rozin, Lowery, Imada, & Haidt, 1999; Russell & Giner-Sorolla, 2013) to investigate whether different emotional expressions lead observers to infer qualitatively different norms. We develop and test the idea that observed angry and disgusted reactions elicit norm inferences that differ in their relative reference to standards of autonomy and purity.

EASI theory postulates that people are influenced by other people's emotional expressions through two different mechanisms: affective reactions to other people's expressions, brought about by processes such as emotional contagion (Hatfield,

**CONTACT** Marc W. Heerdink  [m.w.heerdink@uva.nl](mailto:m.w.heerdink@uva.nl)

 Supplemental data for this article can be accessed here <https://doi.org/10.1080/02699931.2018.1476324>

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Cacioppo, & Rapson, 1994), and contextualised cognitive inferences that people make about the meaning of an emotional expression (Van Kleef, 2009). These cognitive inferences are often based on back-tracking the appraisals that underlie an emotional expression within a given context (De Melo, Carnevale, Read, & Gratch, 2014; Hareli & Hess, 2010; Manstead & Fischer, 2001; Van Kleef, 2016). The resulting social influence of emotional expressions has been demonstrated across a range of contexts, including negotiations (e.g. Sinaceur & Tiedens, 2006; Van Kleef, De Dreu, & Manstead, 2004), attitude change (Calanchini, Moons, & Mackie, 2016; Van Kleef, Van den Berg, & Heerdink, 2015), leadership (Sy, Côté, & Saavedra, 2005; Van Kleef et al., 2009), and majority influence in groups (Heerdink, Van Kleef, Homan, & Fischer, 2013). Here, we test if emotional influence also contributes to the formation and transmission of social norms.

Social norms are group-based representations of what others value and reject, which are embedded in the cognitive group prototype (Cialdini & Goldstein, 2004; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), and may thus be subject to the inferential processes described in EASI theory. Social norms come in two types: Descriptive norms refer to perceptions of what others typically or normally do, whereas injunctive norms refer to how people “ought” to behave, i.e. rules or beliefs about behaviour in relation to moral standards. Given that violations of injunctive norms are inappropriate and immoral (Cialdini et al., 1990), they trigger strong negative emotional reactions in other group members (Cameron, Lindquist, & Gray, 2015; Haidt, 2001; Van Kleef, Wanders, Stamkou, & Homan, 2015). Hence, observed negative emotional reactions to behaviour may be especially informative about the injunctive norms in the situation.

In the moral emotions literature, attempts have been made to understand which negative emotional reactions are triggered by different types of norm violations (e.g. Horberg et al., 2011; Hutcherson & Gross, 2011; Royzman, Atanasov, Landy, Parks, & Gepty, 2014; Rozin et al., 1999; Russell & Giner-Sorolla, 2013). Among the most prominent perspectives in this field is work that builds on an especially well-defined and clear mapping proposed by Rozin et al. (1999). According to this perspective, anger is associated with violations of the standard of autonomy, which refers to concerns about fairness, reciprocity, and the prevention of harm (Graham, Haidt, & Nosek, 2009, 2011; Horberg et al., 2011; Rozin et al.,

1999). Disgust, on the other hand, is proposed to be associated with violations of purity, broadly defined. Which behaviours in particular trigger disgust is subject to debate (for a review, see Russell & Giner-Sorolla, 2013), but disgust is often associated with threats of contagion and violations of the integrity of the body, which are (together with threats to the purity of the mind) encompassed in the standard of purity (Graham et al., 2009, 2011; Horberg et al., 2009; Rozin et al., 1999).<sup>1</sup> Although these associations are conceptually absolute and uncorrelated, empirical work finds a degree of overlap between these associations (e.g. Cameron et al., 2015), and we accordingly consider them as relative rather than absolute. It should furthermore be noted that although this work provided the impetus and background for our research, different perspectives on these relations between elicitors and moral emotions exist (e.g. Cameron et al., 2015; Hutcherson & Gross, 2011; Molho, Tybur, Güler, Balliet, & Hofmann, 2017; Royzman et al., 2014; Russell & Giner-Sorolla, 2013), and we return to these alternative interpretations in the General Discussion.

With regard to their interpersonal functions, angry and disgusted expressions are similar in the sense that they reflect hostility in the expresser (Izard, 1977) and, as a result, increase the social distance between the expresser and the target (e.g. Fischer & Manstead, 2008; Heerdink, Van Kleef, Homan, & Fischer, 2015). The interpersonal consequences of angry expressions have been extensively investigated in a variety of contexts (e.g. Heerdink et al., 2013; Tiedens, 2001; Van Kleef et al., 2009), and initial evidence indicates that angry reactions do indeed signal to observers that a norm was violated (Hareli, Moran-Amir, David, & Hess, 2013), even though this research did not compare the consequences of angry and disgusted reactions. In stark contrast, to the best of our knowledge only two articles have examined the interpersonal consequences of disgust expressions, showing that disgusted expressions communicate moral condemnation (Kupfer & Giner-Sorolla, 2017), and that in a dyadic context, others' expressions of anger and disgust elicit guilt and shame in targets through their associations with standards of autonomy and purity, respectively (Giner-Sorolla & Espinosa, 2011).

Here we build on and extend this previous work by investigating the effects of expressions of anger and disgust from a third-person perspective. That is, rather than examining responses of *targets* of emotional expressions (see Giner-Sorolla & Espinosa,

2011), we examine the inferences third-party observers draw from others' angry and disgust expressions. Understanding such third-party effects is important theoretically, because prior theorising suggests that people can learn the norms and values of a group by witnessing others' emotional reactions to behaviours (Keltner & Haidt, 1999). Because observers are less involved in the situation than targets, and because the number of observers of a single emotional reaction can potentially be great, our third-party perspective helps understand a key mechanism through which standards for behaviour become shared group norms.

### The present research

To summarise, observers are able to reverse-engineer the cause of emotional reactions (De Melo et al., 2014; Hareli & Hess, 2010; Van Kleef et al., 2004), and different types of injunctive norm violations have been proposed to elicit different emotional reactions depending on the moral standard involved (e.g. Horberg et al., 2011; Rozin et al., 1999). Thus, an observer may be able to infer both if and why (in terms of the moral standard) a behaviour is inappropriate from a third person's emotional reaction to that behaviour. Given that emotional influence is most potent in relatively ambiguous situations (Van Kleef, 2016), we presented participants with emotional reactions to behaviour that was ambiguous with regard to whether it violated a norm.

Associations between violations of moral standards and emotional reactions tend not to be absolute; rather, moral violations elicit a combination of different negative emotional reactions, which only differ in their relative strength (Cameron et al., 2015; Landmann & Hess, 2018). For instance, violations of purity standards may elicit both anger and disgust, but the disgusted reaction tends to be stronger. Accordingly, in reversing these associations, we adopt a relative (rather than absolute) conceptualisation of the associations between norm violations and emotional reactions. We expected that, compared to expressions of disgust, expressions of anger result in an injunctive norm inference that is relatively more based in autonomy standards than purity standards. In other words, we expected the difference between autonomy-based and purity-based injunctive norms to be greater (more positive) following an angry expression than following a disgusted expression (i.e. an Expression  $\times$  Norm Type interaction). We conducted

six studies to test this hypothesis, of which we report three studies in full in the present article. These studies test the hypothesis in the context of snacking (Study 1), drinking alcohol (Study 2), and unspecified behaviour (Study 3). The three remaining studies can be found in the online supplement.<sup>2</sup>

### Study 1

In Study 1 we operationalised the main hypothesis in the context of the greedy eating of shared, unhealthy snacks, to which a friend reacted with anger or disgust. It is ambiguous whether this behaviour is a norm violation, and we believed it would be possible both to construe it as a purity violation (threatening the integrity of the body), and as an autonomy violation (unfair to the others and harming their chances of enjoying the snacks). Building on the idea that the informational influence of emotional expressions is most pronounced in ambiguous situations (Adam & Brett, 2015; Van Kleef, De Dreu, & Manstead, 2010), we explored the moderating influence of a minimal amount of accompanying information in the form of a statement by the friend: one referring to personal standards ("I think you cannot do this!"), and one to universal standards ("You cannot do this!"). Although the aforementioned perspectives suggest that the effects should be clearest when there is no accompanying statement, we were unsure whether such minimal information would have an effect and therefore we did not formulate a formal hypothesis.

### Method

#### Participants and design

We used a 2 (Expression: anger or disgust)  $\times$  3 (Statement: none, universal, or personal) between-subjects design. Data collection was stopped after (daily) observation showed that at least 180 participants had participated (which was a requirement for an unrelated study that followed the present one). Twenty-two participants failed the forced-choice Statement manipulation check, and three participants responded with insufficient effort (one responded very quickly and highly repetitively on the Likert scales [ $|z|$  scores for both criteria  $> 2$ ]; two had a negative person-total correlation; Huang, Liu, & Bowling, 2015). Excluding these participants, the final sample consisted of 174 Dutch participants (124 women, 41 men, 9 undisclosed;  $M_{\text{age}} = 20.58$ , range 18–56), who participated in exchange for course credit.

### Vignette

Participants considered the following vignette (emotion word matching expression condition):

A group of female friends goes out to dinner together. They sit down and order some drinks. When the drinks are served, they also receive a platter of fried snacks on the house. As soon as the snacks are on the table, Elsa attacks the platter and starts eating greedily. [Angrily/Disgusted], Hanna watches how Elsa eats one snack after another.

The vignette ended there (no statement condition), or continued with "As Elsa starts on her nth snack, Hanna exclaims", followed by "You cannot do this!" (universal statement condition), or "I think you cannot do this!" (personal statement condition).<sup>3</sup>

### Questionnaire

The vignette was followed by a questionnaire which included four items to measure purity-based norm violations (e.g. "In this group, people who eat many snacks are seen as distasteful") and five items to measure autonomy-based norm violations (e.g. "In this group, Elsa's behaviour is considered selfish"; 1 = does not apply, 7 = applies very well).<sup>4</sup> In line with our literature review above, we formulated the autonomy items in terms of unfairness and harm to the interests of others, and the purity items in terms of threats to the integrity of the body (e.g. risk of contamination). The other measures in the questionnaire were included for exploratory purposes.<sup>5</sup> A two-factor EFA with oblimin rotation indicated that the purity-related items loaded on the first factor (loadings  $\geq .64$ ;  $\alpha = .84$ ) and the autonomy-related items loaded on the second factor (loadings  $\geq .47$ ;  $\alpha = .64$ ). There were no cross-loadings (all  $< .30$ ). Thus, the factor analysis distinguished between autonomy-based and purity-based norm violations.

Later in the questionnaire, participants also indicated to what extent anger, disgust and six other emotions (contempt, happiness, fear, grief, disappointment, and surprise) characterised Hannah's

reaction to Elsa's behaviour (expression manipulation check; 1 = does not apply, 7 = applies very well). A forced-choice read check asked participants which of the three possible statements they had encountered in the story and was used for excluding inattentive participants (see Participants and Design above).

### Results

In this and all following studies, we used a model comparison approach (e.g. Crawley, 2007) for all analyses involving a within-subjects factor (e.g. Norm Type). Using the lme4 package for R, we fit a full model (all possible main effects and interactions), a simple model (all main effects and the Expression  $\times$  Within-subjects factor interaction, which is the hypothesised effect), and a null model (only the within-subjects factor main effect). We report the simplest model that does not sacrifice predictive power compared to the full model (i.e. a model comparison yields a non-significant result at  $\alpha = .05$ ). We mention all significant effects but focus our interpretation on the key effects and estimated effect sizes as the change in total  $R^2$  ( $\Delta R^2$ ; assessed using the Jaeger, Edwards, Das, and Sen [2017] method) resulting from including the term in a model that contains all lower-order and other same-order model terms. Finally, for the sake of clarity, we indicate between-subjects factors with (B) and within-subjects factors (repeated measures) with (W). Table 1 reports descriptive statistics and inter-correlations for the measures discussed below.

### Manipulation checks

Comparison of full, simple, and null models with repeated measures of Perceived Expression (anger = -1, disgust = 1) demonstrated that the full model fit best: full vs. null,  $\chi^2(10) = 97.51$ ,  $p < .001$ ; full vs. simple,  $\chi^2(6) = 27.17$ ,  $p < .001$ . The only significant effects in this model were the Expression (B)  $\times$  Perceived Expression (W) interaction ( $\chi^2[1] = 92.66$ ,  $p < .001$ ,  $\Delta R^2 = 14.9\%$ ), the main effects of Statement

**Table 1.** Descriptive statistics and simple effects for the main dependent measures depending on condition in Study 1.

	No statement		Universal statement		Personal statement		Correlations			
	Anger	Disgust	Anger	Disgust	Anger	Disgust	1	2	3	4
Anger MC (1)	6.07 (0.83)	4.81 (1.20)	5.63 (0.89)	4.84 (1.27)	6.07 (0.77)	5.68 (0.65)	1.00			
Disgust MC (2)	5.04 (1.32)	6.65 (0.75)	4.63 (1.45)	5.72 (1.20)	4.86 (1.51)	5.58 (1.15)	0.02	1.00		
Autonomy NV (3)	5.42 (0.69)	5.18 (0.93)	5.41 (0.39)	4.71 (0.79)	5.54 (0.67)	5.32 (0.62)	0.29*	-0.11	1.00	
Purity NV (4)	3.15 (1.33)	4.28 (1.05)	3.45 (1.35)	3.82 (1.27)	3.79 (1.40)	3.75 (1.25)	-0.04	0.44*	-0.01	1.00

Note: MC = Manipulation check. NV = Norm violation. Reported descriptive statistics are means (SDs).

\* $p < .050$ .

(B) ( $\Delta R^2 = 2.3\%$ ), the two-way Statement (B)  $\times$  Perceived Expression (W) interaction ( $\Delta R^2 = 3.0\%$ ), and the three-way Expression (B)  $\times$  Statement (B)  $\times$  Perceived Expression (W) interaction ( $\Delta R^2 = 2.0\%$ ; all  $ps < .016$ ). The Statement main effect and interactions suggest that, relative to no statement, the universal statement decreased both perceived anger and disgust, whereas the individual statement increased perceived anger and decreased perceived disgust. More importantly, the relative differences between perceived anger and disgust were in the intended direction in all conditions, even though the three-way interaction indicated that this pattern varied somewhat in clarity depending on the statement that was made (Table 1): relatively more anger than disgust was perceived in the angry expression conditions, and relatively more disgust than anger was perceived in the disgust expression conditions. This pattern was most pronounced in the no-statement condition and least pronounced in the personal-statement condition. Because the pattern was as intended across all conditions, we considered the manipulation successful and proceeded with the analyses as planned.

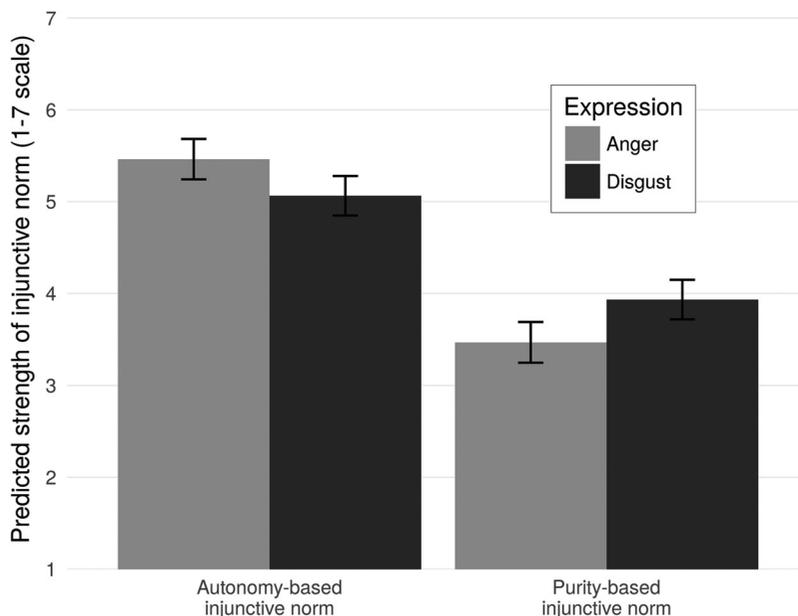
### Norm inferences

We compared full, simple and null models with repeated measures of Norm Type (autonomy-based = -1, purity-based = 1). The simple model fit best: null vs. simple,

$\chi^2(4) = 18.50$ ,  $p = .001$ ; simple vs. full,  $\chi^2(6) = 12.34$ ,  $p = .055$ . The only significant effects in the simple model were the predicted Expression (B)  $\times$  Norm Type (W) interaction ( $\chi^2[1] = 15.55$ ,  $p < .001$ ,  $\Delta R^2 = 2.7\%$ ) and the main effect of Norm Type (W) ( $\chi^2[1] = 208.62$ ,  $p < .001$ ,  $\Delta R^2 = 35.0\%$ ). Although overall, the autonomy-based norm inference was stronger than the purity-based norm inference, the relative differences between conditions were in line with the hypothesis (Figure 1): An angry expression led to a relatively greater (positive) difference between autonomy-based norm inferences and purity-based norm inferences than a disgusted expression. The raw descriptive statistics (Table 1) additionally suggest that the expected pattern was more pronounced in the “no statement” and “universal statement” conditions than in the “personal statement” condition, although this pattern was not strong enough for the full model to fit better.

### Discussion

We found initial support for the idea that observing an angry reaction to an ambiguous behaviour fuels inferences that the behaviour is inappropriate because it is a relatively stronger violation of autonomy standards than purity standards, compared to a disgusted reaction. This general pattern was somewhat qualified by the exploratory manipulation of statement. In



**Figure 1.** Estimated strength of inferred autonomy-based and purity-based injunctive norms depending on emotion expressed about a behaviour (Study 1). Prediction is based on model coefficients; error bars represent 95% confidence intervals.

general, the clearest pattern was obtained when no further statement accompanied the emotional reaction. This finding is relevant in light of the proposition that interpersonal effects of emotional expressions are strongest in ambiguous situations (Van Kleef et al., 2010). These findings suggest that emotional expressions may take on heightened importance as carriers of normative information in more ambiguous situations.

## Study 2

In this study, we tested the main hypothesis with a different behaviour: drinking alcohol in a university building. We used video stimuli instead of vignettes to increase the degree of realism compared to Study 1, which also guards against the possibility that our findings are merely the result of semantic associations between a verbal description of the emotional reaction, and the items measuring the DV.<sup>6</sup> We also aimed to test the generalizability of our findings across different expressers, by varying the expresser's social power. Power entails the capacity to exert influence on others (Keltner, Gruenfeld, & Anderson, 2003) and to define group norms (Taggar & Ellis, 2007), suggesting that emotional expressions of high-power individuals may have a greater impact on observers' norm inferences than emotional expressions of low-power individuals. On the other hand, high-power individuals are more likely to violate group norms (Keltner et al., 2003; Van Kleef, Wanders, Stamkou, & Homan, 2015), and low-power individuals' behaviour may therefore be more diagnostic of actual group norms. In light of these contradictory perspectives, we examined the potential moderating role of the expresser's social power in an exploratory fashion.

## Method

### Participants and design

We used a 3 (Expression: none, anger, or disgust)  $\times$  2 (Power: high or low) between-subjects design. We combined two datasets that were collected in the lab in separate student projects. The first dataset was based on a 2 (Emotion: anger vs disgust)  $\times$  2 (Power: high vs low) between-subjects design ( $N = 96$ ). The second dataset was collected eight months later and had a 2 (Emotion: neutral vs disgust)  $\times$  2 (Power: high vs low) between-subjects design ( $N = 73$ ). The samples were drawn from two subsequent

cohorts of the same population (undergraduates at the local university) and differed neither in gender ( $p = 0.776$ ) nor in age ( $p = 0.229$ ). The sample sizes were determined by the maximum number of participants we were able to recruit in the available time for data collection. One participant responded with insufficient effort (identical responses on 98.6% of all Likert scales in the experiment; Huang et al., 2015) and 14 participants did not refer to alcohol in their description of the situation (i.e. they failed to mention the words "alcohol" or "liquor" [including misspelled versions], particular alcoholic drinks or brands (e.g. "vodka", "bacardi"), or any euphemism referring to alcohol (e.g. "a little something"), suggesting that they had not understood the video. Excluding these participants, the final sample consisted of 154 participants (110 women and 44 men;  $M_{\text{age}} = 22.03$ , range 18–56).

### Video clips

In the video clips, two women (both acting school students) wait outside a classroom for their seminar to begin. One of the two women wears plain clothes and is identified as student "Lisa". The other woman wears formal clothes and is identified as the teacher (high power condition), or wears plain clothes and is identified as a fellow student (low power condition). Lisa checks her watch, takes a bottle of coke and a spirit flask from her bag, pours the contents of the flask into her bottle, and takes a big gulp. The other woman (depending on condition) does keeps a neutral expression, expresses anger, or expresses disgust. This actor was coached to express the appropriate emotions using pictures of prototypical anger and disgust expressions from the Facial Action Coding System (Ekman & Friesen, 1978). Except for the manipulations, the video clips were identical.

### Questionnaire

Following viewing the video clip, participants were first asked to describe the video in their own words (we analysed these responses to exclude participants, see above) and then completed a questionnaire containing the DV measures. A two factor EFA with oblimin rotation on the six norm-related items ( $1 = \text{disagree completely}$ ,  $7 = \text{agree completely}$ ) distinguished between those measuring autonomy-based inappropriateness (3 items,  $\alpha = .60$ ) and those measuring purity-based inappropriateness (3 items,  $\alpha = .79$ ) (see note 4). The power manipulation was checked using 5 items ( $\alpha = .85$ ). An example item is "The [teacher/fellow student] is powerful." ( $1 = \text{disagree completely}$ ,

**Table 2.** Descriptive statistics for the main variables in Study 2 and their zero-order inter-correlations.

	Low Power			High Power			Correlations					
	None	Anger	Disgust	None	Anger	Disgust	1	2	3	4	5	
Anger MC (1)	2.81 (1.80)	5.57 (1.59)	4.55 (1.36)	2.47 (1.50)	5.13 (1.58)	4.20 (1.43)	1.00					
Disgust MC (2)	4.13 (1.67)	6.13 (1.14)	6.73 (0.78)	3.00 (1.94)	4.57 (1.83)	6.63 (0.55)	0.43*	1.00				
Power MC (3)	2.89 (1.12)	3.79 (1.14)	3.54 (1.06)	4.72 (0.99)	5.11 (0.57)	4.76 (1.16)	0.18*	-0.05	1.00			
Autonomy NV (4)	3.65 (1.29)	4.01 (1.07)	4.03 (1.06)	3.71 (1.29)	4.20 (1.07)	3.90 (0.99)	0.27*	0.17*	0.25*	1.00		
Purity NV (5)	3.33 (1.58)	3.55 (1.35)	4.37 (1.41)	3.24 (1.34)	3.26 (1.04)	4.07 (1.58)	0.08	0.30*	0.04	0.29*	1.00	

Note: Reported descriptive statistics are means (*SDs*). MC = Manipulation Check. NV = Norm Violation.

\* $p < .050$ .

7 = *agree completely*). The questionnaire also included several exploratory measures (see note 5). Later in the questionnaire, participants indicated to what extent anger, disgust, and six other emotions (contempt, happiness, fear, grief, pride, and disappointment) characterised the expresser's reaction (emotion manipulation check; 1 = *not at all*, 7 = *very much*).

## Results

Descriptive statistics for the main variables and their inter-correlations are displayed in Table 2. Orthogonal contrast coding was used to compare the conditions: Contrast 1 compared the no expression condition (coded as -1) to the angry and disgusted expression conditions combined (each coded as 0.5) and Contrast 2 compared the angry expression condition (-1) to the disgusted expression condition (1), and was therefore the contrast of focal interest. We report omnibus  $\chi^2$  statistics and  $\Delta R^2$  for the model terms, and betas for the contrasts plus their 95% percentile-based confidence interval obtained with the bootMer method for R (100,000 resamples).

### Manipulation checks and data analysis

Analysis of the manipulation checks indicated that the manipulations were not orthogonal, i.e. each manipulation also affected the manipulation check for the other manipulation. On the power manipulation check, we found a main effect of Power,  $F(1, 150) = 68.05$ ,  $p < .001$ ,  $\eta^2 = 0.31$ , but also an unintended main effect of Expression,  $F(2, 150) = 3.69$ ,  $p = .027$ ,  $\eta^2 = .05$ . Similarly, on the expression manipulation checks we found not only the intended Emotional Expression (B)  $\times$  Perceived Expression (W) interaction ( $\chi^2[2] = 58.45$ ,  $p < .001$ ,  $\Delta R^2 = 6.7\%$ ), but also a main effect of Power ( $\chi^2[1] = 9.59$ ,  $p = .002$ ,  $\Delta R^2 = 2.1\%$ ) and a marginally significant Power (B)  $\times$  Emotional Expression (B)  $\times$  Perceived Expression (W) interaction ( $\chi^2[2] = 5.79$ ,  $p = .055$ ,  $\Delta R^2 = 0.6\%$ ).

This non-orthogonality of our manipulations was circumvented by exchanging the power manipulation for the power manipulation check in all models, while retaining the expression manipulation as a predictor. This estimates the effect of expression while partialling out consequences of perceived differences in power.<sup>7</sup> Re-analysing the expression manipulation checks (perceived anger = -1, perceived disgust = 1) in this fashion, we found that the simple model fit best (simple vs. null:  $\chi^2[5] = 132.41$ ,  $p < .001$ ; simple vs. full:  $\chi^2[5] = 8.91$ ,  $p = .113$ ). The only significant effects in this model were the Perceived Expression (W) main effect ( $\chi^2[1] = 97.04$ ,  $p < .001$ ,  $\Delta R^2 = 11.6\%$ ), the Expression (B) main effect ( $\chi^2[2] = 105.12$ ,  $p < .001$ ,  $\Delta R^2 = 25.5\%$ ), and the Expression (B)  $\times$  Perceived Expression (W) interaction ( $\chi^2[2] = 57.42$ ,  $p < .001$ ,  $\Delta R^2 = 6.8\%$ ). Overall, participants perceived more disgust than anger,  $\beta = 0.279$  [0.207, 0.351]. Furthermore, Contrast 1 had a significant main effect ( $\beta = 0.819$  [0.658, 0.980]), but did not interact with Perceived Expression ( $\beta = 0.041$  [-0.070, 0.153]). Contrast 2 did not have a significant main effect ( $\beta = 0.049$  [-0.064, 0.162]), but did interact with Perceived Expression in the intended direction ( $\beta = 0.300$  [0.221, 0.379]). This indicates that, overall, an expression increased ratings of both disgust and anger, and relative differences between perceived anger and perceived disgust corresponded with the manipulated expressions: An angry expression led to a greater difference between perceived anger and perceived disgust than a disgusted expression. Replacing the power manipulation by the power manipulation check thus effectively resolved the non-orthogonality of the manipulation, allowing us to proceed with our hypothesis testing.

### Norm inferences

We compared null, full, and simple models of the norm inferences with repeated measure of Norm Type (autonomy-based = -1, purity-based = 1). The simple model fit best: simple vs. null,  $\chi^2[6] = 27.37$ ,  $p < .001$ ; simple vs. full,  $\chi^2[5] = 4.69$ ,  $p = .455$ . The only

significant effects in the simple model were the Expression (B) main effect ( $\chi^2[2] = 9.20, p = .010, \Delta R^2 = 3.5\%$ ) and the Expression (B)  $\times$  Norm Type (W) interaction ( $\chi^2[2] = 12.66, p = .002, \Delta R^2 = 2.7\%$ ; Figure 2).<sup>8</sup> Contrast 1 had a near-significant main effect across both norm inferences,  $\beta = 0.202$  [−0.001, 0.404], but did not interact with Norm Type,  $\beta = 0.044$  [−0.106, 0.249]. Contrast 2 had a significant main effect across both norm inferences,  $\beta = 0.149$  [0.005, 0.291], and, more importantly, interacted with Norm Type,  $\beta = 0.186$  [0.079, 0.293]. The coefficients indicate that perceptions of both types of norm violation inferences were higher when a negative emotion was expressed, and especially so if disgust was expressed. Moreover, and in line with the hypothesis, an angry reaction led to a relatively greater difference between perceived autonomy-based and purity-based norm inferences than a disgusted reaction.

### Discussion

In Study 2, we found support for our main hypothesis in a different setting, with a different focal behaviour, and using non-verbal stimuli. Compared to a neutral reaction, both negative emotional reactions triggered inferences that a norm had been violated, with the disgusted reaction being more potent than the angry reaction. More importantly, the nature of the inferred norm differed depending on whether anger or disgust was expressed. In keeping with our hypothesis, compared to expressions of disgust, expressions of anger shifted the inferred norm to be relatively more based in autonomy standards than in purity standards.

### Study 3

In Study 3, we aimed to gauge the generality of the hypothesised effect by eliminating any contextual information that may steer norm inferences in one direction or another. For this purpose, we used a vignette that contained only the bare minimum information about the people involved, and left the target behaviour unspecified. Three additional methodological improvements were made: First, we included the Moral Foundations Questionnaire (MFQ; Graham et al., 2011) in addition to the scale used in the previous studies to cross-validate our measure of norm inferences. Second, reflecting methodological developments during the period in which these studies were run, we determined our sample size

using a power analysis. And finally, we included a more global injunctive norm violation measure to test our conceptualisation of autonomy-based and purity-based norms as subtypes of injunctive norms.

## Method

### Participants and design

We used a three conditions between-subjects design (Expression: none, anger, or disgust). A simulation-based power analysis suggested that we needed around 400 complete cases to achieve a power of 0.80 for an interaction that explains 3% of the variance (Study 1). We oversampled by 10% to allow exclusion of insufficient effort responding, and paid 436 international CrowdFlower workers \$0.70 to complete the online survey. We excluded 37 participants with a negative person-total correlation, or both a repetitive pattern of responding (>90% same or highly similar responses given sequentially) and being exceptionally fast or slow (>2 SDs faster or slower than the mean log-averaged completion time). 399 complete cases were available for analysis (233 women, 165 men, 1 undisclosed;  $M_{\text{age}} = 36.75$ , range 18–76).

### Vignette

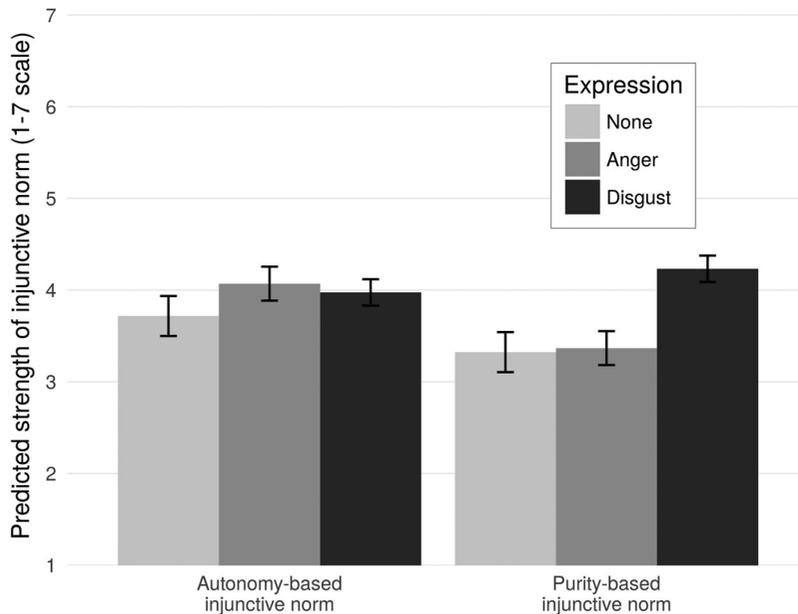
Participants imagined the following scenario:

Walking to work, you pass by a company where two employees are outside, talking. You don't know either of them, but we will call them Ben and Tom for the remainder of this study. Ben and Tom are apparently talking about something that Ben did, earlier this morning.

Depending on condition, the vignette ended there (no expression condition) or stated “You notice that Tom is expressing strong [anger / disgust] about Ben's behaviour.” (anger and disgust conditions, respectively).

### Questionnaire

Two nine-item questionnaires were used to measure norm inferences (all answered on 1 = *does not apply at all*, 5 = *applies very well*). The first was based on the one used in the previous studies and adapted to fit the current scenario: Three items measured autonomy-based norm inferences and two items measured purity-based norm inferences (see note 4). Four additional items measured the inappropriateness of the behaviour more generally to tap into the overall injunctive norm (e.g. “Ben's earlier behaviour was inappropriate”,  $\alpha = .88$ ). The second questionnaire consisted



**Figure 2.** Estimated strength of inferred autonomy-based and purity-based injunctive norms depending on emotion expressed about a behaviour (Study 2). Prediction is based on model coefficients; error bars represent 95% confidence intervals.

of the autonomy-related subscales (harm and fairness) and the purity subscale from the MFQ (Graham et al., 2011). MFQ items were rephrased to refer to Ben's behaviour, and we changed the item labelled DISGUSTING in the purity scale to "Ben did something repulsive" to avoid a direct reference to disgust. After preliminary analysis, we dropped the two reverse-scored MFQ items because of near-zero item-total correlations with the other MFQ items in their respective subscales. To simplify analysis and presentation, we combined the autonomy-related items from both questionnaires to form an autonomy-based norm inference scale (8 items,  $\alpha = .93$ ) and the purity-related items to form a purity-based norm inference scale (4 items,  $\alpha = .88$ ). Note that combining items from the two questionnaires resulted in highly reliable scales, which suggests that the MFQ and our own norm violation items overlap considerably in terms of the construct that is measured. Analyses using either only our own scales, or only the MFQ scales yielded identical conclusions as those reported below. In addition, the results are similar across the MFQ harm and fairness subscales. We therefore only report analyses on the combined scales.

Later in the questionnaire, participants also indicated to what extent anger, disgust, and three more emotions had been expressed by Tom (1 = not at all, 5 = extremely).

## Results

Descriptive statistics and zero-order correlations for the main measures can be found in Table 3. We report the same orthogonal contrasts and the same statistics as in Study 2.

### Manipulation checks

Comparing full and null models of the expression manipulation checks (perceived anger = -1; perceived disgust = 1), we found that the full model fit best,  $\chi^2(4) = 298.90$ ,  $p < .001$ . The only significant effects in this model were the Emotional Expression (B) main effect ( $\chi^2[2] = 287.62$ ,  $p < .001$ ,  $\Delta R^2 = 32.6$ ) and the Emotional Expression (B)  $\times$  Perceived Expression (W) interaction ( $\chi^2[2] = 91.42$ ,  $p < .001$ ,  $\Delta R^2 = 4.1\%$ ). The coefficients indicated that Contrast 1 had a main effect ( $\beta = 0.806$  [0.714, 0.899]), but did not interact with Perceived Expression ( $\beta = -0.028$  [-0.087, 0.031]). Contrast 2, on the other hand, had no main effect ( $\beta = 0.003$  [-0.078, 0.084]), but did interact with Perceived Expression ( $\beta = 0.248$  [0.197, 0.299]). Together, these coefficients indicate that both perceived anger and perceived disgust were higher in the two emotional conditions compared to the no expression condition. More importantly, the difference between perceived anger and perceived disgust was greater in the angry

**Table 3.** Means, SDs (in brackets) and zero-order correlations of the main measures in Study 3.

	Expression			Correlations				
	None	Anger	Disgust	1	2	3	4	5
Anger MC (1)	2.27 (1.16)	4.27 (0.86)	3.62 (1.13)	1.00				
Disgust MC (2)	2.26 (1.15)	3.48 (1.00)	4.15 (1.06)	0.56*	1.00			
Autonomy NV (3)	2.13 (1.01)	3.15 (0.70)	3.20 (0.74)	0.58*	0.58*	1.00		
Purity NV (4)	2.10 (1.02)	3.07 (0.78)	3.39 (0.83)	0.54*	0.61*	0.87*	1.00	
Inappropriateness (5)	2.30 (1.04)	3.27 (0.76)	3.40 (0.88)	0.57*	0.56*	0.86*	0.87*	1.00

Note: MC = Manipulation check. NV = Norm violation. The reported inter-correlations between dependent measures are generally large because the no expression condition led to much lower scores on all dependent variables than both emotional expression conditions.

\* $p < .050$ .

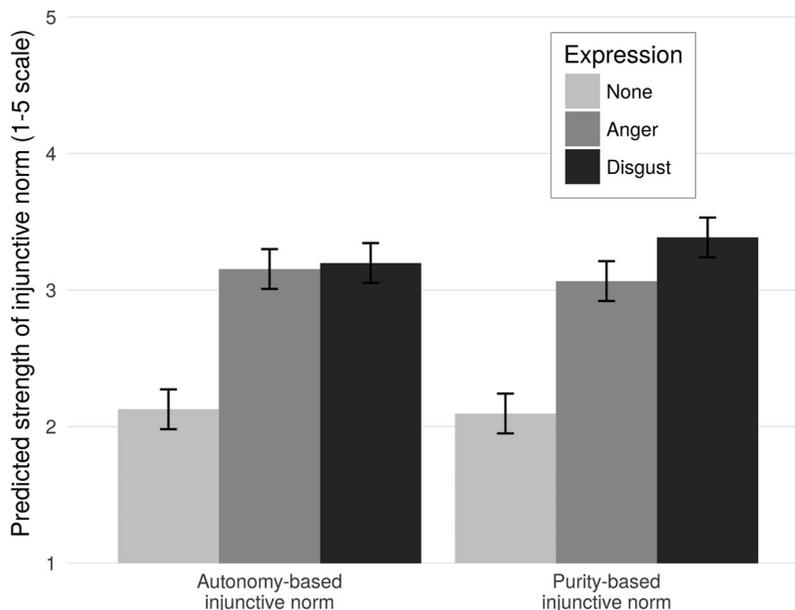
condition than in the disgust condition. Thus, the expression manipulation was successful.

### Norm inferences

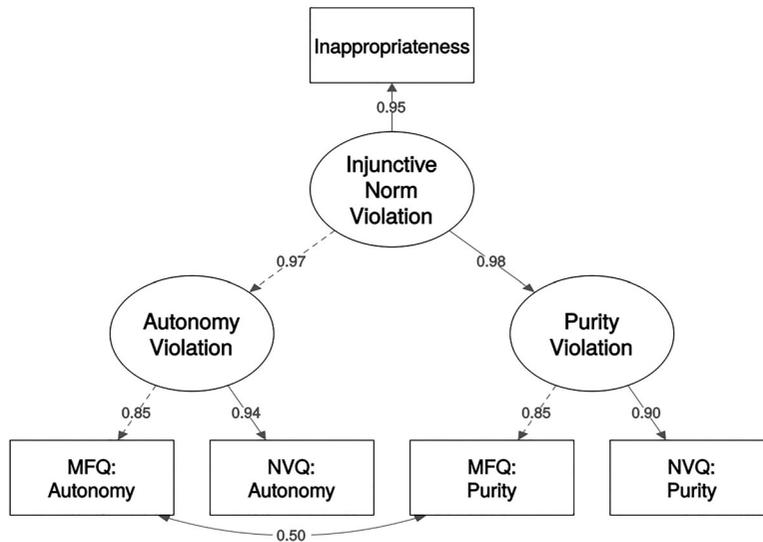
We compared null and full models of the norm inferences with repeated measures of Norm Type (autonomy-based = -1, purity-based = 1). The full model fit better than the null model,  $\chi^2(4) = 158.11$ ,  $p < .001$ . The only significant effects in this model were the Emotional Expression (B) main effect ( $\chi^2[2] = 161.97$ ,  $p < .001$ ,  $\Delta R^2 = 27.0\%$ ) and the Emotional Expression (B)  $\times$  Norm Type (W) interaction ( $\chi^2[2] = 22.79$ ,  $p < .001$ ,  $\Delta R^2 = 0.4\%$ ; Figure 3). The coefficients indicated that Contrast 1 had a main effect ( $\beta = 0.727$  [0.613, 0.840]), but did not interact with Perceived Expression ( $\beta = 0.027$  [-0.008, 0.061]). Contrast 2, on the other hand, had no main effect

( $\beta = 0.091$  [-0.007, 0.188]), but did interact with Perceived Expression ( $\beta = 0.069$  [0.039, 0.099]). Thus, both norm inferences were higher in the emotional conditions than in the no expression condition. Moreover, and in line with our hypothesis, the difference between autonomy-based and purity-based norm inferences was greater in the angry expression condition than in the disgusted expression condition.

On the general inappropriateness measure, we also found a main effect of Emotional Expression ( $F(2, 369) = 59.79$ ,  $p < .001$ ,  $\eta^2 = 0.23$ ). The coefficient for Contrast 1 was significant ( $\beta = 0.69$ ,  $t = 10.87$ ,  $p < .001$ ), but the coefficient for Contrast 2 was not ( $\beta = 0.07$ ,  $t = 1.23$ ,  $p = .220$ ), indicating that both emotional reactions increased perceptions that a norm was violated to a similar degree.



**Figure 3.** Estimated strength of inferred autonomy-based and purity-based injunctive norms depending on emotion expressed about a behaviour (Study 3). Prediction is based on model coefficients; error bars represent 95% confidence intervals.



**Figure 4.** Confirmatory Factor Analysis in which moral violations are subtypes of injunctive norm violations (standardized coefficients) for Study 3. MFQ: Moral Foundations Questionnaire (Graham et al., 2011). NVQ: Norm violation questionnaires as developed for this article. Model fit statistics:  $\chi^2(2) = 1.02$ ,  $\chi^2/df = 0.51$ , CFI = 1.00, TLI = 1.01, RMSEA = 0.00.

### Relation between moral standards and global injunctive norm

A final question we aimed to answer in this study is to what extent the data support our construal of standards of autonomy and purity as subtypes of injunctive norms. We conducted a Confirmatory Factor Analysis (DWLS estimator, mean-and-variance adjusted fit statistics) using the lavaan package for R to test this idea. The model is displayed in Figure 4: We used the MFQ and norm violation subscales as separate indicators of purity violations and autonomy violations. “Injunctive norm violation” was defined as a superordinate latent construct, as an explanation for the observed inter-correlations between purity violations, autonomy violations, and judged inappropriateness. The only parameter that we added was a correlation between the two MFQ scales, purity and autonomy, which was necessary to obtain a good fit to the data. The model fit the data well,  $\chi^2(2) = 1.02$ ,  $\chi^2/df = 0.51$ , CFI = 1.00, TLI = 1.01, RMSEA = 0.00. Thus, the results support a construal of moral standards as underlying injunctive norms.

### Discussion

In Study 3, we used a more abstract situation to study the effects of emotional expressions on norm inferences. We found that angry and disgusted expressions did not differ in their effectiveness for signalling

general inappropriateness, but once again found that angry expressions led to relatively stronger autonomy-based norm inferences than purity-based norm inferences, compared to disgusted expressions. Moreover, we found that the autonomy-based and purity-based norm inference measures from the previous studies formed reliable scales when combined with items from the validated MFQ (Graham et al., 2011), which enhances confidence in the validity of the measure from Studies 1 and 2. Finally, a confirmatory factor analysis yielded support for our conceptualisation of injunctive norms as a superordinate construct that encompasses different moral standards.

### General discussion

Drawing on EASI theory (Van Kleef, 2016) and the literature on moral emotions (e.g. Horberg et al., 2011; Rozin et al., 1999; Russell & Giner-Sorolla, 2013), we investigated which injunctive norms observers infer from another person’s emotional reactions to particular behaviour. We predicted that observing angry reactions to behaviour leads to qualitatively different injunctive norm inferences than observing disgusted reactions to the same behaviour, with angry reactions shifting the injunctive norm inference to be relatively more based in the moral standard of autonomy than purity compared to disgusted reactions. Three studies, employing different operationalisations of

emotional expression and different types of behaviours, found support for this prediction. We conclude that observers not only infer from angry and disgusted reactions *that* an injunctive norm was violated, but that the emotional reaction also shapes inferences of *why* the behaviour is inappropriate, in terms of the moral standard that was violated.

### **Contribution and theoretical implications**

This research demonstrates that emotional expressions help transmit social norms, especially to newcomers and people unfamiliar with the group (i.e. socialisation; Levine & Moreland, 1994). Prior research on the role of observed behaviour for norm learning indicates that information about what others typically do, such as statistical information and the behaviour of relevant others, is used to make norm inferences (e.g. Cialdini et al., 1990; Paluck & Shepherd, 2012). Emotional expressions may function as a complementary source of information as they help observers infer what is appropriate in a situation (the injunctive norm; see also Hareli et al., 2013) rather than what is typical (the descriptive norm). Our results indicate that angry and disgusted expressions do not only convey general inappropriateness, but also convey additional information about *why* a particular behaviour is seen as inappropriate, namely because it violates different moral standards to different degrees.

As mentioned in the introduction, there are alternative perspectives on the relation between moral emotions and moral violations (e.g. Cameron et al., 2015; Chapman & Anderson, 2013; Hutcherson & Gross, 2011; Royzman et al., 2014). One of these perspectives propose primacy of anger over disgust (Royzman et al., 2014), whereas another suggest primacy of disgust over anger (Hutcherson & Gross, 2011; Tybur, Lieberman, Kurzban, & DeScioli, 2013) in moral judgment. Our findings suggest that, in the interpersonal domain, there is no primacy of either expression. Although Study 2 found that disgusted (versus angry) expressions increased both types of inferred norms in the context of alcohol ingestion, Study 3 did not replicate this pattern in the context of unspecified behaviour, and also revealed no differential impact of angry and disgusted expressions on general inferences of inappropriateness. A potentially relevant contextual factor is the extent to which a behaviour harms the expresser (e.g. Chapman & Anderson, 2013; Kupfer & Giner-Sorolla, 2017; Molho

et al., 2017; Young & Tsoi, 2013). If the harm is primarily to oneself (versus to others), as with the alcohol ingestion in Study 2, anger may signal general immorality more potently than disgust.

Another area of discussion concerns the degree to which disgust is related to non-bodily purity (or divinity) violations (e.g. Chapman & Anderson, 2013; Hutcherson & Gross, 2011; Royzman et al., 2014; Russell & Giner-Sorolla, 2013). Given that our measure of purity violations focused only on body-related elicitors of disgust (e.g. distastefulness), as is common in measures of purity in related work (e.g. Clifford, Iyengar, Cabeza, & Sinnott-Armstrong, 2015; Graham et al., 2011), our findings do not provide a direct answer to this question. However, Study 3 indicated disgusted expressions and angry expressions were equally effective in conveying a general sense of inappropriateness, and that both inferred (bodily) purity violations and autonomy violations are related to judgments of inappropriateness. This suggests that even if disgusted expressions convey only that a behaviour constitutes a bodily purity violation, this might be sufficient to moralise the behaviour more broadly. Alternatively, given the inherent ambiguity in perceiving the causes of others' emotional expressions, it is also possible that without further information, perceivers make both bodily and non-bodily purity inferences at the same time.

Finally, constructionist perspectives challenge the notion of specificity implied by the theory discussed in the Introduction, and shifts the focus more toward what is shared between emotions and morality: core affect and conceptual activation (Cameron et al., 2015). In support of this perspective, we indeed found that perceived anger and perceived disgust were correlated in our studies, as were and autonomy-based and purity-based norm violations. However, neither shared core affect, nor shared conceptual activation can readily explain our findings. An overlap in core affect is an unlikely explanation because the person who apparently experienced the emotion (the expresser) was a different person than the person making the moral inference (the perceiver). Conceptual activation is an unlikely explanation because we found similar results using both verbal (Studies 1 and 3) and non-verbal stimuli (Study 2), despite the former being intuitively more conceptual than the latter—a finding that resonates with a growing corpus of research indicating that emotions have similar social effects regardless of whether they are expressed via the face, the voice, the body, or in

words (Van Kleef, 2017). Future research could further examine shared “ingredients” that may explain why certain types of moral violations and emotional expressions are relatively more likely to be associated than others on the interpersonal level.

More generally, our findings show that the literature on the morality-related functions of disgust (e.g. Rozin et al., 1999; Tybur et al., 2013) provides a useful basis for developing an understanding of the interpersonal functions of disgust expressions. In contrast to the extensive research on the interpersonal consequences of anger (e.g. Heerdink et al., 2013; Sinaceur & Tiedens, 2006; Tiedens, 2001; Van Kleef et al., 2004), research has largely ignored the consequences of disgust at the interpersonal level (for exceptions, see Giner-Sorolla & Espinosa, 2011; Kupfer & Giner-Sorolla, 2017). Future research may fruitfully build on the insight that behaviour that elicits disgusted expressions shifts norm inferences toward purity standards, which may moralise the behaviour such that both the behaviour itself, and the people enacting the behaviour are avoided by observers (see also Molho et al., 2017).

### **Limitations and future directions**

The effects reported here are of small to medium magnitude (a meta-analysis on all studies, including those in the online supplement, estimated the difference in relative strength of autonomy-based and purity-based norms between the angry and disgusted expression conditions at  $d = 0.43$ , 95% CI: 0.23 to 0.64; see the online supplement for more details), and the effects appear to be attenuated when more contextual information is provided (e.g. adding a verbal utterance as in Study 1, or using less ambiguous situations as in the supplemental studies). We consider this boundary condition for our effects a demonstration of the idea that emotional expressions are most informative when they help disambiguate a situation, which is a tenet of EASI theory (Van Kleef, 2016) that has seldom been tested directly (for an exception, see Adam & Brett, 2015). The observed modest effect size may thus be primarily due to our decision to isolate the effects of facial emotional expressions. In real life, facial emotional expressions are likely to co-occur with congruent information, such as a specific tone of voice, bodily posture, and supporting statements. Such additional information is likely to (a) reduce confusion about the emotion that is expressed and (b) reduce degrees of freedom for inferring the

meaning of the expression in the observer, thereby enhancing the effects that we discovered here.

A key future direction for the current work is to investigate how differences in the degree to which perceived norms are based in autonomy or purity standards impact behaviour. We see two potential avenues for such research. First, research might explore how evaluations of new social targets showing the same behaviour are impacted by prior normative information. Do we perceive purity-based norm violators as more “impure” and therefore worthier of avoidance than autonomy-based norm violators (cf., Molho et al., 2017; Tybur et al., 2013)? Second, the likelihood of the observer enacting the inappropriate behaviour might be differentially impacted by specific contextual moderators. Applying the idea that autonomy standards function primarily to prevent harm to others, and purity standards to prevent harm to oneself (Young & Tsoi, 2013), behaviour that elicited an angry reaction may be avoided only in situations where others are present, whereas a disgusted reaction may lead to more universal avoidance of this behaviour. Similarly, taking freely from an abundant resource (e.g. eating snacks, or taking fish from a pond) is less harmful and unfair to others (i.e. autonomy-violating) than taking from a limited resource. Thus, the likelihood of enacting behaviour that elicits angry reactions may be more sensitive to contextual variation in abundance than behaviour that elicits a disgusted reaction (see also Russell & Giner-Sorolla, 2011).

To conclude, we have shown that observers infer qualitatively different injunctive norms from emotional expressions of anger and disgust, and that the differences between these inferred norms correspond to the moral bases associated with these emotions (Horberg et al., 2011; Rozin et al., 1999). This demonstrates that emotional expressions are relevant to the transmission of social norms and socialisation of group members more generally, and contributes to the moralisation of certain behaviours in groups. As acculturation researchers are beginning to show (e.g. Witkamp, Dotsch, Van Halen, Tong, & Wigboldus, 2013), a readiness to attend to emotional cues and a proper understanding of the information contained in emotional expressions could therefore be a key asset for newcomers who strive for acceptance in a social group.

### **Notes**

1. In this framework, a third moral emotion, contempt, is associated with violations of community standards,

- concerned with the duty to uphold obligations to the community or its hierarchy (Rozin et al., 1999). The present work focuses only on anger and disgust for two reasons: First, facial expressions of anger and disgust are more easily recognised by observers than expressions of contempt (Russell, 1994; Wagner, 2000), which reduces contempt's potential for interpersonal influence compared to anger and disgust. Second, the morality literature primarily contrasts anger with disgust (e.g. Cameron et al., 2015; Russell & Giner-Sorolla, 2013) and the distinctiveness of contempt as an emotion has been questioned (Prinz, 2007), which complicates the formulation of hypotheses about contempt.
- The division of studies over main text and supplement reflects a trade-off between completeness and information value. The three studies reported in the main text are the strongest methodologically, and therefore the most informative. The studies in the supplement generally used 'stronger' situations in the sense that they situated the emotional expressions in a context that was more familiar to the participants, which may have interfered with the manipulations and caused a more ambiguous pattern of results.
  - The vignette is translated from Dutch to English. The original Dutch vignette may be obtained from the first author.
  - All items used to measure the dependent variables are included in the supplement.
  - In Studies 1 and 2, the remaining items in the main questionnaire (28 and 29 items, respectively) were intended to develop scales for other standards for behaviour, including inferences about body and health ideals, formal rules and prohibitions, the descriptive norm, and anticipated consequences in terms of formal sanctions or exclusion, and to explore how these inferences are impacted by emotional influence. All items were worded to fit the context of the study, and all items were completed in fully randomised order. Initial analyses indicated that the factorial structure of these items was difficult to interpret and inconsistent across studies, suggesting limited validity of these items, and exploratory analyses yielded inconsistent findings. We therefore discontinued these explorations in Study 3.
  - We thank an anonymous reviewer for this interpretation.
  - Given that angry expressions have been found to be perceived as signs of power (Tiedens, 2001), the analytic procedure reported in the main text also potentially partials out effects of angry expressions on norm inferences that are mediated by power perceptions, resulting in a slightly more conservative test. Ignoring the non-orthogonality of the manipulations in the analyses yields virtually identical results in the main analysis: The simple model fits best (null versus simple:  $\chi^2(5) = 22.28, p < .001$ , simple versus full:  $\chi^2(5) = 1.84, p = 0.871$ ), with only a significant main effect of Expression ( $\chi^2(2) = 9.38, p = .009$ ) and a significant Expression  $\times$  Norm Type interaction ( $\chi^2(2) = 12.91, p = .002$ ). Contrast 1 has neither a significant main effect ( $\beta = 0.24 [-0.08, 0.56]$ ) nor interacts with Norm Type ( $\beta = 0.11 [-0.27, 0.50]$ ), whereas Contrast 2 has no significant main effect ( $\beta = -0.07 [-0.30, 0.15]$ ), but does interact with Norm Type ( $\beta = 0.48 [0.21, 0.76]$ ), in line with our predictions ( $R = 25,000$ ).
  - The coefficient of the power manipulation check was marginally significant ( $\chi^2[1] = 3.82, p = .051$ ): Both injunctive norm inferences were somewhat stronger to the extent that the actor was seen as more powerful,  $\beta = 0.160 [0.000, 0.319]$ .

## Acknowledgments

We thank Karlijn Nagtegaal for her assistance in creating the videos and helping set up Study 2.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

Part of this work was supported by a grant from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek [grant number NWO 452-09-010] awarded to the last author.

## ORCID

Marc W. Heerdink  <http://orcid.org/0000-0002-6931-6519>  
Gerben A. van Kleef  <http://orcid.org/0000-0003-0823-7654>

## References

- Adam, H., & Brett, J. M. (2015). Context matters: The social effects of anger in cooperative, balanced, and competitive negotiation situations. *Journal of Experimental Social Psychology, 61*, 44–58. doi:10.1016/j.jesp.2015.07.001
- Calanchini, J., Moons, W. G., & Mackie, D. M. (2016). Angry expressions induce extensive processing of persuasive appeals. *Journal of Experimental Social Psychology, 64*, 88–98. doi:10.1016/j.jesp.2016.02.004
- Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review, 19*, 371–394. doi:10.1177/1088868314566683
- Chapman, H. A., & Anderson, A. K. (2013). Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological Bulletin, 139*, 300–327. doi:10.1037/a0030964
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591–621. doi:10.1146/annurev.psych.55.090902.142015
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*, 1015–1026. doi:10.1037/0022-3514.58.6.1015
- Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods, 47*, 1178–1198. doi:10.3758/s13428-014-0551-2
- Crawley, M. J. (2007). *The R book* (1st ed.). Chichester, England, United Kingdom: Wiley.

- De Melo, C. M., Carnevale, P. J., Read, S. J., & Gratch, J. (2014). Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology, 106*, 73–88. doi:10.1037/a0034251
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Fischer, A. H., & Manstead, A. S. R. (2008). Social functions of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 456–468). New York, NY: The Guilford Press.
- Giner-Sorolla, R., & Espinosa, P. (2011). Social cuing of guilt by anger and of shame by disgust. *Psychological Science, 22*, 49–53. doi:10.1177/0956797610392925
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029–1046. doi:10.1037/a0015141
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366–385. doi:10.1037/a0021847
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834. doi:10.1037/0033-295X.108.4.814
- Hareli, S., & Hess, U. (2010). What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition & Emotion, 24*, 128–140. doi:10.1080/02699930802613828
- Hareli, S., Moran-Amir, O., David, S., & Hess, U. (2013). Emotions as signals of normative conduct. *Cognition & Emotion, 27*, 1395–1404. doi:10.1080/02699931.2013.791615
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York, NY: Cambridge University Press.
- Heerdink, M. W., Van Kleef, G. A., Homan, A. C., & Fischer, A. H. (2013). On the social influence of emotions in groups: Interpersonal effects of anger and happiness on conformity versus deviance. *Journal of Personality and Social Psychology, 105*, 262–284. doi:10.1037/a0033362
- Heerdink, M. W., Van Kleef, G. A., Homan, A. C., & Fischer, A. H. (2015). Emotional expressions as social signals of rejection and acceptance: Evidence from the affect misattribution paradigm. *Journal of Experimental Social Psychology, 56*, 60–68. doi:10.1016/j.jesp.2014.09.004
- Horberg, E. J., Oveis, C., & Keltner, D. (2011). Emotions as moral amplifiers: An appraisal tendency approach to the influences of distinct emotions upon moral judgment. *Emotion Review, 3*, 237–244. doi:10.1177/1754073911402384
- Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology, 100*, 828–845. doi:10.1037/a0038510
- 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*, 719–737. doi:10.1037/a0022408
- Izard, C. E. (1977). *Human emotions*. Boston, MA: Springer US. Retrieved from <http://link.springer.com/10.1007/978-1-4899-2209-0>
- Jaeger, B. C., Edwards, L. J., Das, K., & Sen, P. K. (2017). An  $R^2$  statistic for fixed effects in the generalized linear mixed model. *Journal of Applied Statistics, 44*, 1086–1105. doi:10.1080/02664763.2016.1193725
- Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*, 265–284. doi:10.1037/0033-295X.110.2.265
- Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion, 13*, 505–521. doi:10.1080/026999399379168
- Kupfer, T. R., & Giner-Sorolla, R. (2017). Communicating moral motives: The social signaling function of disgust. *Social Psychological and Personality Science, 8*, 632–640. doi:10.1177/1948550616679236
- Landmann, H., & Hess, U. (2018). Testing moral foundation theory: Are specific moral emotions elicited by specific moral transgressions? *Journal of Moral Education, 47*, 34–47. doi:10.1080/03057240.2017.1350569
- Levine, J. M., & Moreland, R. L. (1994). Group socialization: Theory and research. *European Review of Social Psychology, 5*, 305–336. doi:10.1080/14792779543000093
- Manstead, A. S. R., & Fischer, A. H. (2001). Social appraisal. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 221–232). New York, NY: Oxford University Press.
- 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*, 609–619. doi:10.1177/0956797617692000
- Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology, 96*, 574–587. doi:10.1037/a0011989
- Paluck, E. L., & Shepherd, H. (2012). The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network. *Journal of Personality and Social Psychology, 103*, 899–915. doi:10.1037/a0030015
- Prinz, J. J. (2007). *The emotional construction of morals*. New York, NY: Oxford University Press.
- Royzman, E., Atanasov, P., Landy, J. F., Parks, A., & Gepty, A. (2014). CAD or MAD? Anger (not disgust) as the predominant response to pathogen-free violations of the divinity code. *Emotion, 14*, 892–907. doi:10.1037/a0036829
- Rozin, P., Lowery, L., Imada, L., & 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. doi:10.1037/0022-3514.76.4.574
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102–141. doi:10.1037/0033-2909.115.1.102
- Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science, 2*, 360–364. doi:10.1177/1948550610391678
- Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: What it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological Bulletin, 139*, 328–351. doi:10.1037/a0029319
- Sinaceur, M., & Tiedens, L. Z. (2006). Get mad and get more than even: When and why anger expression is effective in negotiations. *Journal of Experimental Social Psychology, 42*, 314–322. doi:10.1016/j.jesp.2005.05.002

- Sy, T., Côté, S., & Saavedra, R. (2005). The contagious leader: Impact of the leader's mood on the mood of group members, group affective tone, and group processes. *Journal of Applied Psychology, 90*, 295–305. doi:10.1037/0021-9010.90.2.295
- Taggar, S., & Ellis, R. (2007). The role of leaders in shaping formal team norms. *The Leadership Quarterly, 18*, 105–120. doi:10.1016/j.leaqua.2007.01.002
- Tiedens, L. Z. (2001). Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral. *Journal of Personality and Social Psychology, 80*, 86–94. doi:10.1037/0022-3514.80.1.86
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: Blackwell.
- Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and structure. *Psychological Review, 120*, 65–84. doi:10.1037/a0030778
- Van Kleef, G. A. (2009). How emotions regulate social life: The emotions As social information (EASI) model. *Current Directions in Psychological Science, 18*, 184–188. doi:10.1111/j.1467-8721.2009.01633.x
- Van Kleef, G. A. (2016). *The interpersonal dynamics of emotion: Toward an integrative theory of emotions as social information*. Cambridge: Cambridge University Press.
- Van Kleef, G. A. (2017). The social effects of emotions are functionally equivalent across expressive modalities. *Psychological Inquiry, 28*, 211–216. doi:10.1080/1047840X.2017.1338102
- Van Kleef, G. A., De Dreu, C. K. W., & Manstead, A. S. R. (2004). The interpersonal effects of emotions in negotiations: A motivated information processing approach. *Journal of Personality and Social Psychology, 87*, 510–528. doi:10.1037/0022-3514.87.4.510
- Van Kleef, G. A., De Dreu, C. K. W., & Manstead, A. S. R. (2010). An interpersonal approach to emotion in social decision making: The emotions as social information model. *Advances in Experimental Social Psychology, 42*, 45–96. doi:10.1016/S0065-2601(10)42002-X
- Van Kleef, G. A., Homan, A. C., Beersma, B., van Knippenberg, D., van Knippenberg, B., & Damen, F. (2009). Searing sentiment or cold calculation? The effects of leader emotional displays on team performance depend on follower epistemic motivation. *Academy of Management Journal, 52*, 562–580. doi:10.5465/amj.2009.41331253
- Van Kleef, G. A., Van den Berg, H., & Heerdink, M. W. (2015). The persuasive power of emotions: Effects of emotional expressions on attitude formation and change. *Journal of Applied Psychology, 100*, 1124–1142. doi:10.1037/apl0000003
- Van Kleef, G. A., Van Doorn, E. A., Heerdink, M. W., & Koning, L. F. (2011). Emotion is for influence. *European Review of Social Psychology, 22*, 114–163. doi:10.1080/10463283.2011.627192
- Van Kleef, G. A., Wanders, F., Stamkou, E., & Homan, A. C. (2015). The social dynamics of breaking the rules: Antecedents and consequences of norm-violating behavior. *Current Opinion in Psychology, 6*, 25–31. doi:10.1016/j.copsyc.2015.03.013
- Wagner, H. L. (2000). The accessibility of the term “contempt” and the meaning of the unilateral lip curl. *Cognition & Emotion, 14*, 689–710. doi:10.1080/02699930050117675
- Witkamp, B., Dotsch, R., Van Halen, C., Tong, E., & Wigboldus, D. H. J. (2013). *A cross-cultural study on how immigrants' understanding of approval and disapproval relates to successful adaptation (Unpublished Master's thesis)*.
- Young, L., & Tsoi, L. (2013). When mental states matter, when they don't, and what that means for morality: When mental states matter for morality. *Social and Personality Psychology Compass, 7*, 585–604. doi:10.1111/spc3.12044