Interpersonal Consequences of Deceptive Expressions of Sadness

Christopher A. Gunderson1, Alysha Baker2, Alona D. Pence3, and Leanne ten Brinke4

Abstract
Emotional expressions evoke predictable responses from observers; displays of sadness are commonly met with sympathy and help from others. Accordingly, people may be motivated to feign emotions to elicit a desired response. In the absence of suspicion, we predicted that emotional and behavioral responses to genuine (vs. deceptive) expressers would be guided by empirically valid cues of sadness authenticity. Consistent with this hypothesis, untrained observers (total N = 1,300) reported less sympathy and offered less help to deceptive (vs. genuine) expressers of sadness. This effect was replicated using both posed, low-stakes, laboratory-created stimuli, and spontaneous, real, high-stakes emotional appeals to the public. Furthermore, lens models suggest that sympathy reactions were guided by difficult-to-fake facial actions associated with sadness. Results suggest that naive observers use empirically valid cues to deception to coordinate social interactions, providing novel evidence that people are sensitive to subtle cues to deception.

Keywords
emotion, facial expressions, sympathy, deception

Received December 15, 2020; revision accepted October 26, 2021

Emotions have multiple components—subjective, physiological, and behavioral—that serve as functional reactions to stimuli in our environment (e.g., Keltner & Gross, 1999). Among those behavioral reactions are emotional facial expressions that communicate one’s emotional state and activate predictable responses from others (e.g., Keltner & Haidt, 1999; Van Kleef, 2016). Accordingly, people may be motivated to feign emotional facial expressions to evoke a desired response from others (DePaulo, 1992; Goffman, 1959). These attempts at manipulation are thought to be highly successful; research consistently finds that explicit lie detection accuracy is at or near chance (Bond & DePaulo, 2006), and that observers perform similarly poorly in discriminating genuine versus deceptive expressions of emotion (Ekman & O’Sullivan, 1991; Hess & Kleck, 1994; Porter & ten Brinke, 2008; Stel & van Dijk, 2018; Zloteanu, 2019). Here, we bring together the social functional account of emotions (Keltner & Gross, 1999; Van Kleef, 2009) and findings on behavioral cues to deceptive emotional expressions (Ekman, 2003; Ekman et al., 1988; Frank et al., 1993; Hill & Craig, 2002; Hurley & Frank, 2011; Porter et al., 2012; Porter & ten Brinke, 2008; ten Brinke et al., 2012; ten Brinke & Porter, 2012) to suggest that—despite poor explicit accuracy—observers will experience different emotional and behavioral reactions to genuine versus deceptive displays of sadness, indicating sensitivity to subtle cues of deception on the face.

Emotional Facial Expressions
There is considerable debate about the nature of emotions and their expression, including the extent to which emotional facial expressions accurately reflect inner emotional experience, the heterogeneity or variation in the expression of each emotion, and the universality of emotional facial expressions (e.g., Barrett et al., 2019; Crivelli & Fridlund, 2019; Keltner et al., 2019; Russell, 1995). Despite the continuing controversy, it is clear that observers infer emotion in patterns of facial action. When presented with an expression that includes a furrowed brow and downturned lip corners, people tend to infer sadness. Granted, the reliability of this inference is highest when participants are forced to choose from a short list of emotion labels and lowest when permitted to label freely (e.g., Barrett et al., 2019; Calvo & Nummenmaa, 2016; Elfenbein & Ambady, 2002; Srinivasan & Martinez, 2019).

1University of Denver, CO, USA
2Okanagan College, Kelowna, British Columbia, Canada
3Lewis & Clark College, Portland, OR, USA
4The University of British Columbia, Kelowna, Canada

Corresponding Author:
Leanne ten Brinke, The University of British Columbia, Okanagan Campus, 1147 Research Road, Kelowna, British Columbia, Canada V1V 1V7.
Email: leanne.tenbrinke@ubc.ca
Social Functions of Emotions

From a social functional perspective, expressions of emotion provide information to social partners and evoke specific responses from them (e.g., Keltner & Gross, 1999; Niedenthal & Brauer, 2012; Shariff & Tracy, 2011; see also Crivelli & Fridlund, 2018). Humans infer emotion in the facial expressions of others and experience inferential, affective, and behavioral responses to those expressions, coordinating social interactions (Van Kleef, 2009). Consistent with the social functional account of emotion, expressions of distress may lead to behavioral reactions by observers aimed at reducing the expresser’s negative emotional state and contributing to social cohesion. For example, the experience of sadness is defined as resulting from a perceived or actual loss, particularly when one feels unable to prevent or restore that loss (Levenson, 1999; Smith & Lazarus, 1993). Accordingly, expressions of sadness—compared with other emotions—are perceived as useful when attempting to elicit help from others (Hackenbracht & Tamir, 2010). In response to expressions of sadness, observers generally experience sympathy and concern (Batson & Shaw, 1991; Eisenberg et al., 1989; Marsh & Ambady, 2007). In addition to these emotional responses, observers are likely to offer help to individuals expressing sadness. For instance, Small and Verrochi (2009) found that expressers who appeared sad (vs. happy or neutral) elicited greater sympathy and more charitable donations from observers.

Thus, regardless of the correlation between one’s facial expression and their inner emotional experience, it is clear that facial expressions serve a social and communicative function (Fridlund, 2014; Keltner et al., 2006; Keltner & Gross, 1999; Lee & Anderson, 2017; Matsumoto et al., 2008; Niedenthal & Brauer, 2012; Van Kleef, 2009). Indeed, it should not be assumed that all expressions reflect a genuine emotional experience (e.g., Hess & Kleck, 1990). Expressions can be deliberately adopted in an act of self-presentation aimed at setting the tone and influencing the direction of a social interaction (Goffman, 1959). Specifically, simulations occur when an expression is adopted but unaccompanied by a felt emotion (e.g., a smile without feeling happiness), and masks conceal a felt emotion with a different, unfelt expression. Both simulated and masked expressions are considered deceptive if they are expressed intentionally to mislead others (Vrij, 2008).

Deceptive facial expressions, however, appear to differ in subtle ways from genuine ones. Specifically, genuine and deceptive displays differ on the degree of “prototypical completeness” (e.g., Ekman et al., 1990; Porter et al., 2012), “leakage” of inconsistent emotional expressions (e.g., Ekman et al., 1988; Porter et al., 2012; Porter & ten Brinke, 2008), symmetry (Dopson et al., 1984; Ekman et al., 1981; Hill & Craig, 2002), and in the timing of onset and offset of the expression (Ekman & Friesen, 1982; Hess & Kleck, 1990; see also Horic-Asselin et al., 2020). These differences are more likely to appear in high- versus low-intensity emotional contexts (Porter et al., 2012). In addition, in extremely high-stakes acts of deception in which individuals plead for the return of a missing relative who they have in fact killed, facial actions differ across liars and truth tellers (ten Brinke & Porter, 2012; see also Wright Whelan et al., 2015). Specifically, deceptive pleaders are less likely to express prototypical signs of sadness, including furrowed brows and downturned lip corners, relative to genuine pleaders. Although emotional cues to deception will likely be highly context dependent, it appears that deceptive expressions of sadness can be accompanied by subtle but perceivable cues.

Emotional and Behavioral Reactions to Deceptive Emotional Displays

To the extent that interpersonal reactions predicted by the social functional perspective of emotion (Keltner & Gross, 1999; Van Kleef, 2009) are sensitive to cues to deception in facial expression, reactions to deceptive (vs. genuine) expressions are likely to differ. Indeed, research reveals that deceptive (vs. genuine) expressions of happiness during hotel check-in and restaurant customer service interactions resulted in lower perceived friendliness, satisfaction, and lower customer–employee rapport (Grande et al., 2005). In addition, in the context of a mock fundraiser, observers donated less money to expressers who simulated (vs. genuinely) expressed happiness (Hidieg & van Kleef, 2017). Thus, although observers are generally not suspicious of deception in social interactions (i.e., Ostrich effect; Vrij, 2008; see also Levine, 2014), their emotional and behavioral reactions may still differ in response to them.

Regardless of one’s motives for feigning an emotion (e.g., garner rapport by appearing happy, or sympathy by appearing sad), research has yet to illustrate the process by which these emotional reactions occur in response to genuine versus deceptive emotional expressions. We argue that reactions to deceptive (vs. genuine) expressions differ because deceptive expressions are likely to be incomplete—missing prototypical facial actions that are difficult to activate voluntarily (i.e., reliable action units [AUs]; Mehu et al., 2012). Specifically, in sadness, the combination of AUs 1 and 4—which bring the inner eyebrows upward and together—is difficult to deliberately engage (Ekman et al., 1980; Gosselin et al., 2010; Rosenberg & Ekman, 1994; see also Namba et al., 2017). Consequently, the incomplete appearance of feigned (vs. genuine) sadness expressions may result in less sympathy and help. We argue that observers are attuned to the activation of such reliable AUs and that attention to these cues underlies their affective and behavioral responses to deceptive (vs. genuine) expressers, even in the absence of suspicion of deceit.

The Current Research

In Studies 1a and 1b, we test the hypotheses that observers will report more sympathy (Hypothesis 1 [H1]) and donate
more money (Hypothesis 2 [H2]) to expressers who display genuine (vs. deceptive) facial expressions of sadness, using highly controlled laboratory stimuli. In Study 2, we tested H1 and H2 using a sample of real, high-stakes emotional lies and truths. Specifically, we studied observers’ emotional and (hypothetical) financial reactions to genuine relatives and deceptive murderers pleading for the return of their missing loved one. We further hypothesized, using a lens model approach (Brunswik, 1956), that ecologically valid cues of expressers’ emotional facial actions would inform observers’ sympathy reactions (Hypothesis 3 [H3]). We expected that the activation of reliable AUs associated with the prototypical display of sadness (AU 1 + 4) would be difficult to fake and therefore an empirically valid cue to genuine sadness, which would elicit sympathy and donations from observers (Gosselin et al., 2010; Mehu et al., 2012; ten Brinke & Porter, 2012). All research was IRB-approved, sample sizes were determined a priori, and we report all data exclusions (if any), manipulations, and measures (Simmons et al., 2011). No studies in this article were preregistered.

**Study 1a**

The goal of Study 1a was to test H1 to H3, using a sample of highly controlled, laboratory-created expressions of genuine and deceptive sadness.

**Method**

**Participants.** An a priori power analysis suggested that a sample size of 210 would be required to detect small-to-medium effect of $d = .25$ in a one-factor (veracity: genuine, deceptive) within-subjects design, with $1-\beta = .95$ and $\alpha = .05$ (two-tailed). Two hundred fifty-one participants were recruited through Amazon Mechanical Turk (MTurk) and they participated in exchange for US$1.75.¹ This sample included 93 females and 158 males with a mean age of 35.16 years ($SD = 9.70$). This study was only available to MTurk workers who were living in the United States and who had approval ratings of 90% or higher on their previous work.

**Materials**

*Expresser stimuli.* This study used a total of 80 videos as stimuli ($M_{length} = 6.43$ s)—two from each of 40 expressers ($M_{age} = 21.51, SD = 4.79, n = 34$ females). Videos were created in a previous study (Porter & ten Brinke, 2008) wherein expressers viewed a slideshow of images from the International Affective Picture System (IAPS; Lang et al., 1999) and show a front, close-up view of expressers’ faces. Expressers were instructed to display emotional facial expressions in response to images they were shown. For this study, only videos of expressers displaying genuine and deceptive sadness were included; *genuine sadness* videos depicted expressers who were instructed to express sadness in response to an image reliably rated as evoking feelings of sadness, whereas *deceptive (masked) sadness* videos depicted expressers instructed to express sadness in response to an image that was reliably rated as evoking feelings of happiness (Mikels et al., 2005). Although expressions were posed, rather than spontaneous, *genuine* expressions were consistent with the emotion elicited by the image being viewed, whereas *deceptive* expressions misled the observer as to the expressers’ actual emotional state, consistent with the definition of deception offered by Vrij (2008). Each expresser provided one genuine and one deceptive sadness video; however, participants in this study (i.e., observers) never saw the same expresser twice. Because expressers provided both genuine and deceptive videos to our overall stimuli set, extraneous factors that might affect observer reactions (e.g., gender, attractiveness) were held constant across the genuine and deceptive stimuli, providing a highly controlled set of stimuli in which to test our hypotheses.

**Coding expressers’ facial actions.** In the context of a previous study (Porter & ten Brinke, 2008), frame-by-frame coding for the presence and duration of prototypical expressions of each of the universal emotional expressions (i.e., happiness, sadness, fear, disgust, anger, and surprise; Ekman & Friesen, 1986; Ekman et al., 1987) was completed. This coding scheme was based on prototypical expressions described in the Emotional Facial Action Coding System (FACS; Ekman et al., 2002), but the upper and lower regions of the face were coded independently as cues to deception are often more subtle than full-face expressions (Porter et al., 2012; Porter & ten Brinke, 2008). As cues to deception may be subtle and some actions are more involuntary than others (Ekman, 2003; Rinn, 1984; ten Brinke & Porter, 2012), this level of detail is necessary to quantify the subtlety of facial cues to deception. The upper face region comprises the eyes and forehead, whereas the lower face region is defined as the nose, mouth, cheek, and chin. Specifically, the expression of sadness in the upper face region involves raising and pulling together the inner eyebrows, which is accomplished by activating the inner *frontalis* and *corrugator supercili*. Sadness, expressed in the lower face region, involves pulling the lip corners down and raising the chin to form a pout, which is accomplished by activating the *depressor anguli oris* and the *mentalis*. Coders examined videos frame by frame, noting onset and offset times of emotional expressions in the upper face and separately in the lower face. These onset and offset times were used to calculate duration scores. Of particular interest in this investigation are duration scores for the prototypical expression of sadness, in the upper and lower face, for each expresser video. Specifically, the duration of facial actions, consistent with expressions of sadness, were summed to create duration scores for the upper and lower face, separately.

As a manipulation check for this study, emotional facial expression coding data from Porter and ten Brinke (2008) were reanalyzed. For the upper face, genuine expressions

¹ Assuming a Cohen’s $d$ of .25, a two-tailed test with $\alpha = .05$, and a desired power of $1-\beta = .95$, a sample size of 210 would be required.
(M = 1.96, SD = 2.03) were consistent with prototypical displays of sadness for a longer duration compared with deceptive (masked) expressions (M = 1.18, SD = 1.70), t(38) = 2.56, p = .01, 95% confidence interval (CI) M\_diff = [.17, 1.40], d = .41, 95% CI = [.08, .74]. For the lower face, however, genuine expressions (M = 1.09, SD = 1.66) did not differ from deceptive expressions (M = 1.41, SD = 1.74), t(38) = −1.23, p = .23, 95% CI M\_diff = [−.85, .21], d = −.20, 95% CI = [−.51, .12].

Procedure. Participants were told that they would be watching a series of short videos of students ostensibly reacting to their Graduate Record Exam (GRE) scores. Specifically, they were told that these students had agreed to participate in our research, which involved not reviewing their GRE scores until they came to our lab where we would videotape their reactions. In addition, participants were told that the GRE is difficult and that some students do well, whereas others do not. To increase believability, participants were told that they had been “randomly assigned” to a series of videos of students receiving “disappointing” scores. In reality, there was no other condition of students receiving encouraging scores. Participants watched each video, one at a time; videos were presented without sound and in random order.

In a within-subjects design, participants viewed 20 (10 genuine, 10 deceptive) randomly selected expresser videos. For each expresser, participants were randomly assigned to see either their genuine or deceptive sadness video (i.e., each participant saw 20 unique expressers, half of which were genuine). After each video, participants were asked to rate whether they felt sympathy for the student in the video on a scale of 1 (definitely not) to 7 (definitely yes). Next, participants were told that students can retake the GRE but that it is costly (US$160). Participants were then asked to indicate how much money they would hypothetically donate to the student, on a scale of US$0 to US$160, to help them retake their exam. Two composite variables were created for sympathy ratings: one averaged participants’ sympathy ratings for genuine expressers and the other averaged participants’ sympathy ratings for the deceptive expressers. Similarly, two composite variables were created for hypothetical donations: one averaged participants’ donations to genuine expressers and the other averaged participants’ donations to deceptive expressers. Participants then completed a demographic questionnaire (gender, age, and ethnicity) and reported whether they had any technical difficulties. On average, the study took 15 min to complete.

Finally, participants completed a funnel debriefing to assess suspicion. A research assistant, blind to hypotheses, coded responses to identify any participants who were suspicious. Consistent with the truth-default theory (Levine, 2014), only a small number of participants were identified as being suspicious that some expressers were displaying false emotions (n = 9; 3.6% of the overall sample) or of the cover story (n = 10; 4% of the overall sample). Given the small proportion of suspicious participants, all participants were included in analyses; excluding suspicious participants yielded similar results. A sensitivity power analysis indicated that, with n = 251 participants, we had 80% power to detect an effect as small as d = .18 in a paired-samples t-test with α = .05 (two-tailed). Data and analysis code for all studies presented in this article can be found at https://osf.io/ymznd/

Results

Emotional response to genuine versus deceptive sadness: Sympathy. Consistent with H1, a paired-samples t-test revealed that observers reported more sympathy for expressers displaying genuine (M = 3.90, SD = 1.42), compared with deceptive, sadness (M = 3.81, SD = 1.43), t(250) = 2.31, p = .021, 95% CI M\_diff = [.014,.180], d = .15, 95% CI = [.02,.27].

Emotional response to genuine versus deceptive sadness: Intentions to help. To determine whether observers’ intent to help differed for genuine compared with deceptive expressers, the donation amount given to each type of expresser was examined. A paired-samples t-test revealed that observers donated marginally more money to expressers displaying genuine sadness (M = US$35.56, SD = US$40.28) compared with expressers displaying deceptive sadness (M = US$33.95, SD = US$39.07), t(249) = 1.93, p = .054, 95% CI M\_diff = [−.03,.24], d = .12, 95% CI = [−.002,.247].

Brunswikian lens models: Ecologically valid cues, cue utilization, and achievement. To test H3, we used a Brunswikian lens model to examine whether facial actions associated with prototypical displays of sadness were associated with both expresser veracity and observers’ sympathy responses (Brunswik, 1956; Reynolds & Gifford, 2001). This approach has been used to understand the cues that underlie actual and perceived cues to interpersonal deception (Hartwig & Bond, 2011), facial expressions of pain (Ruben & Hall, 2016), and impressions of psychopathic traits (ten Brinke et al., 2017). Using the lens model nomenclature in the context of this study, facial actions that differentiate genuine from deceptive expressers would be considered ecologically valid cues to veracity. The extent to which observers’ sympathy reactions are influenced by facial actions is referred to as cue utilization. Accuracy represents the extent to which facial actions are both ecologically valid and inform observer sympathy reactions. Specifically, we predicted that the activation of difficult-to-fake (i.e., reliable) facial AUs associated with the prototypical display of sadness in the upper face (i.e., upper face consistency; Ekman et al., 1980; Gosselin et al., 2010; Rosenberg & Ekman, 1994) would be both an ecologically valid cue and one utilized by observers to guide their sympathy reactions, representing accuracy.

Sympathy. A series of Pearson correlations revealed that both upper face sadness, r(37) = .49, 95% CI = [.21,.70],
p = .002, and lower face sadness, r(37) = .55, 95% CI = [0.28, 0.74], p < .001, were associated with expresser veracity. Observers’ sympathy was positively associated with upper face sadness, r(76) = .35, 95% CI = [0.13, 0.53], p = .002, but not lower face sadness, r(76) = .16, 95% CI = [−0.07, 0.37], p = .171. Achievement (i.e., the correlation between expresser veracity and observers’ sympathy) was r(78) = .09, 95% CI = [−0.30, 0.13] p = .432 (see Figure 1).

**Donations.** A series of Pearson correlations revealed that both upper face sadness, r(37) = .49, 95% CI = [0.21, 0.70], p = .002, and lower face sadness, r(37) = .55, 95% CI = [0.28, 0.74], p < .001, were associated with expresser veracity. Observers’ donations, however, were neither associated with upper face sadness, r(76) = .22, 95% CI = [−0.04, 0.42], p = .054, nor with lower face sadness, r(76) = .07, 95% CI = [−0.16, 0.29], p = .56. Achievement (i.e., the correlation between expresser veracity and observers’ donations) was r(78) = .13, 95% CI = [−0.10, 0.34] p = .264 (see Figure 2).

**Study 1b**

In Study 1b, we sought to replicate these effects, using a larger sample of participants viewing a greater number of stimuli.

**Method**

**Participants.** Based on the smallest effect size found in Study 1a (d = .12), an a priori power analysis suggested that a sample size of 431 would be required to detect this effect in a one-factor (veracity: genuine, deceptive) within-subjects design, with 1-β = 0.80 and α = 0.05 (one-tailed). Five hundred participants were recruited through MTurk and they participated in exchange for US$3.00. This sample included 257 males (51.4%), 240 females (48%), and three as “other” gender (0.6%), with a mean age of 39.93 years (SD = 12.05). As in Study 1a, only MTurk workers who were living in the United States and who had approval ratings of 90% or higher on their previous work were accepted.

**Procedure.** Participants completed the same procedure as in Study 1a, except that they viewed a random sample of 30 out of 40 expressers (i.e., 10 more expressers than Study 1a). On average, the study took 16 min to complete.

Similar to Study 1a, funnel debriefing responses revealed that only a small number of participants were suspicious that some expressers were displaying false emotions (n = 35; 7% of the overall sample) or were suspicious of the cover story (n = 10; 2% of the overall sample). Given the small proportion of suspicious participants, these participants were included in analyses. Excluding these participants did not change the results reported below.

Given concerns about the quality of MTurk data (e.g., Dreyfuss, 2018), we also attempted to ensure data quality by including a captcha verification at the beginning of the study and coded responses to the debriefing items. Two (0.4%) participants did not provide sensible answers to debriefing items and were therefore excluded from analyses. Including these participants did not impact the outcome of the analyses reported below. The final sample consisted of four hundred ninety-eight participants. A sensitivity power analysis indicated that, with n = 498 participants, we had 80% power to detect an effect as small as d = .13 in a paired-samples t-test with α = .05 (two-tailed).
Results

Emotional response to genuine versus deceptive sadness: Sympathy. Consistent with Studies 1a and H1, a paired-samples t-test revealed that observers reported more sympathy for expressers displaying genuine ($M = 3.98, SD = 1.28$) versus deceptive sadness ($M = 3.92, SD = 1.28$), $t(497) = 2.71, p = .007, 95\% CI M_{diff} = [0.017, 0.105], d = .12, 95\% CI = [.03,.21]$.

Emotional response to genuine versus deceptive sadness: Intentions to help. Results of a paired-samples t-test revealed that observers donated more money to expressers displaying genuine sadness ($M = US$40.63, $SD = US$41.82) compared with expressers displaying deceptive sadness ($M = US$39.28, $SD = US$40.44), $t(496) = 2.63, p = .009, 95\% CI M_{diff} = [.32, 2.22], d = .12, 95\% CI = [.03,.21]$, supporting H2.

Brunswikian lens models. Similar to Study 1a, we examined the process by which expresser veracity affected observer reactions, using lens models.

Sympathy. A series of Pearson correlations revealed that both upper face sadness, $r(37) = .49, 95\% CI = [.21,.70], p = .002$, and lower face sadness, $r(37) = .55, 95\% CI = [.28,.74], p < .001$, were associated with expresser veracity. Observers’ sympathy was positively associated with upper face sadness, $r(76) = .34, 95\% CI = [.12,.52], p = .003$, but not with lower face sadness, $r(76) = .18, 95\% CI = [.05,.39], p = .117$. Achievement (i.e., the correlation between expresser veracity and observers’ sympathy) was $r(78) = -.02, 95\% CI = [-.24,.20] p = .858$ (see Figure 3).

Donations. A series of Pearson correlations revealed that both upper face sadness, $r(37) = .49, 95\% CI = [.21,.70], p = .002$, and lower face sadness, $r(37) = .55, 95\% CI = [.28,.74], p < .001$, were associated with expresser veracity. Observers’ donations were associated with upper face sadness, $r(76) = .29, 95\% CI = [.07,.48], p = .010$, but not with lower face sadness, $r(76) = .16, 95\% CI = [.06,.37], p = .152$. Achievement (i.e., the correlation between expresser veracity and observers’ donations) was $r(78) = -.01, 95\% CI = [-.23,.21] p = .915$ (see Figure 4).

Study 2

Although Study 1 provided results consistent with our hypotheses, effect sizes were small (Cohen, 1988). It has previously been argued that laboratory-created stimuli may not include robust cues to deception and therefore may not provide observers with strong or reliable signals to guide their emotional and behavioral reactions (e.g., Gunderson & ten Brinke, 2019; von Hippel & Trivers, 2011). Moreover, Study 1 stimuli were posed expressions and greater ecological validity would come from a study of spontaneous (genuine and deceptive) expressions. Accordingly, in Study 2, we sought to replicate Study 1 results and test H1 and H2, using a sample of real, high-stakes, emotional truths and lies. Specifically, we used videos of individuals genuinely or deceptively pleading publicly for the safe return of a missing loved one, which are known to include robust emotional facial cues to deception (ten Brinke & Porter, 2012; Wright Whelan et al., 2015). In addition, we predicted that the activation of reliable facial AUs that serve to bring the inner eyebrows upward and together (AU 1 + 4; Ekman et al., 1980; Gosselin et al., 2010; Rosenberg & Ekman, 1994) would be
both an ecologically valid cue and one used by observers to guide their sympathy reactions and hypothetical donation behavior, testing H3.

**Method**

**Participants.** An a priori power analysis suggested that a sample size of 272 would be required to detect small-to-medium effect of $d_{z} = .20$, with $1-\beta = .95$ and $\alpha = .05$ (one-tailed). Two hundred ninety-eight participants were recruited through MTurk and they completed the study in exchange for a compensation of US$1.75. The sample included 142 females and 156 males, with a mean age of 36.47 years ($SD = 10.92$ years). The study was only available to MTurk workers who were living in the United States and who had approval ratings of 90% or higher on their previous work.

**Materials**

**Pledger stimuli.** A sample of 24 videos was taken from a larger set of videos ($N = 79$) previously collected by ten Brinke and Porter (2012). The videos were chosen using stratified random sampling to ensure an equal distribution of gender and veracity. On average, each video was of 30 s duration ($SD = 33.10$ s, range = 6–153 s). Each video depicts either a relative or friend making a televised plea for the safe return of, or assistance in finding, a missing loved one. Deceptive pleader videos depicted expressers who were later convicted of murdering the missing loved one based on strong physical evidence such as possession of the murder weapon or DNA evidence (see ten Brinke & Porter, 2012, for additional details). Thus, deceptive pledgers asked for help to find a person whose location they knew. Furthermore, deceptive pledgers knew that the missing person would not hear their message and would not return home safely. Genuine pleader videos depicted expressers who were not involved in the disappearance of the missing person; the missing person was either found in the absence of foul play or another individual was found responsible for the person’s disappearance, based on strong physical evidence. Admittedly, genuine and deceptive pledgers differ in more ways than veracity; most salient is that deceptive pledgers have participated in murder, whereas the genuine pledgers have not. Although this may be viewed as a limitation, we view this as an important aspect of the ecological validity afforded by these stimuli (Levine, 2018). In reality, people do not often lie just because they are told to do so by an experimenter (as in Study 1), but instead choose to lie for some purpose that might include concealing past behavior (Bond et al., 2013; Levine et al., 2010; Markowitz & Levine, 2020; Park et al., 2018; ten Brinke et al., 2019). In short, the strengths and limitations of pleader stimuli in Study 2 complement the strengths and limitations of the stimuli employed in Study 1.

**Coding pledgers’ facial actions.** The AUs of interest were coded by a certified FACS coder who had previously demonstrated high reliability, with a second, blind coder when coding briefer portions of these same pleader videos ( Interrater duration scores were highly correlated, $r = .66$–.98; ten Brinke et al., 2012). Specifically, AUs of interest to this study described activation of muscles involved in the expression of prototypical sadness: *frontalis*, *corrugator supercilii*, and *depressor anguli oris* (AUs 1, 4, and 15; Waller et al., 2008). Upper face AUs (1, 4, and 1 + 4) that bring the inner eyebrows upward and together, respectively, were each coded separately from lower face AU (15), which pull the lip corners downward. Coders (blind to veracity) examined videos frame by frame without sound, noting onset and offset times of AUs in the upper face and again for AUs in the lower face. These onset and offset times were used to calculate durations of each AU or combination of AUs. As each pleader video differed in length, data indicate the proportion (duration of expression/total duration) of the plea that each AU (or combination) was present (see ten Brinke et al., 2012).

**Procedure.** Participants were told that they would be watching several videos depicting individuals being interviewed about the disappearance of a missing loved one and would be rating their emotional reaction to each. Importantly, participants were not informed that some of the pledgers were guilty of killing the person they were pleading to find and were, therefore, providing a deceptive appeal. Participants viewed a stratified random sample of 12 out of 24 videos: six genuine pledgers (three males, three females) and six deceptive pledgers (three males, three females). After watching each video (in random order), participants were asked to indicate whether they felt sympathy for each pleader on a scale of 1 (definitely not) to 7 (definitely yes). After participants completed all of their sympathy ratings for each video, participants were shown neutral still images of each of the pledgers they saw and were asked to complete a measure of their intention to help. Participants were told that GoFundMe is a crowdsourced fundraising platform where individuals can donate money to those in need. Participants then indicated how much (hypothetical) money they would donate to each of the pledgers on a scale of US$0 to US$100, using a sliding scale. Similar to Studies 1a and 1b, composite variables were created by averaging participants’ rating of sympathy: one for genuine pledgers and one for deceptive pledgers. Similarly, composite variables were created by averaging participants’ donations to genuine and deceptive pledgers, respectively. Participants then completed demographic questions, reported any technical difficulties, and a funnel debriefing questionnaire. On average, the study took 19 min to complete.

Only a small number of participants recognized any of the pledgers ($n = 12$; 4.0% of entire sample) or were suspicious that some of the pledgers may have been lying ($n = 19$; 6.4% of entire sample). Given the small proportion of suspicious participants, all participants were included in the analyses; excluding these participants did not change the
results reported below. A sensitivity power analysis indicated that, with \( n = 298 \) participants, we had 80% power to detect an effect as small as \( dz = .16 \), with \( \alpha = .05 \) (two-tailed) in a one-factor (veracity: genuine, deceptive) within-subjects design.

**Results**

*Emotional response to genuine versus deceptive pleaders: Sympathy.* A paired-samples \( t \)-test revealed that observers reported more sympathy for genuine (\( M = 5.23, SD = 1.17 \)) compared with deceptive (\( M = 4.90, SD = 1.20 \)) pleaders, \( t(297) = 8.05, p < .001, 95\% CI M_{diff} = [.25, .41], d = .47, 95\% CI = [.35, .59] \).

*Emotional response to genuine versus deceptive pleaders: Intentions to help.* Results of a paired-samples \( t \)-test revealed that observers donated more money to genuine (\( M = \$36.60, SD = 31.53 \)) compared with deceptive (\( M = \$32.36, SD = 29.96 \)) pleaders, \( t(297) = 5.98, p < .001, 95\% CI M_{diff} = [2.85, 5.64], d = .35, 95\% CI = [.23, .46] \).

**Brunswikian lens model: Ecologically valid cues, cue utilization, and achievement**

*Sympathy.* A series of Pearson correlations revealed that AU 1 + 4, \( r(22) = -.48, 95\% CI = [-.74, −.09], p = .018 \), and AU 15, \( r(22) = -.45, 95\% CI = [-.72, −.05], p = .029 \), were associated with pleader veracity (genuine pleaders coded as 0 and deceptive pleaders coded as 1). Observers’ sympathy was positively associated with AU 1 + 4, \( r(22) = .55, 95\% CI = [.19, .78], p = .005 \). Achievement (i.e., the correlation between pleader veracity and observers’ sympathy) was \( r(22) = -.24, 95\% CI = [-.59, .18], p = .254 \) (see Figure 5, see Note 3).

*Donations.* A series of Pearson correlations revealed that AU 1 + 4, \( r(22) = -.48, 95\% CI = [-.74, −.09], p = .018 \), and AU 15, \( r(22) = -.45, 95\% CI = [-.72, −.05], p = .029 \), were associated with pleader veracity. Observers’ sympathy was positively associated with AU 1 + 4, \( r(22) = .55, 95\% CI = [.18, .78], p = .006 \). Achievement (i.e., the correlation between pleader veracity and observers’ donations) was \( r(22) = -.25, 95\% CI = [-.59, .17], p = .243 \) (see Figure 6).

**Discussion**

In a series of studies, observers report experiencing more sympathy and intentions to help genuine versus deceptive expressers of sadness. Although effects were small in the context of posed, low-stakes expressions created in the laboratory (Studies 1a and 1b), effect sizes were larger in the context of spontaneous, high-stakes emotional pleas (Study 2). This highlights the importance of studying ecologically valid stimuli as they may differ from what is created in the lab. Although experimental control is lost in lieu of ecological validity, and the act of deception is likely to become inextricably linked to the behavior it is intended to conceal, we believe that these stimuli better approximate the lies that people actually encounter in their lives (Bond et al., 2013; Levine et al., 2010; Markowitz & Levine, 2020; Park et al., 2018). Collectively, results suggest that observers’ emotional reactions and intentions to help are informed by empirically valid cues of sadness authenticity.

Specifically, lens models suggested that observers’ sympathy and donations in response to genuine versus deceptive
pleaders were informed by the presence of difficult-to-fake facial actions that are associated with prototypical displays of sadness (Gosselin et al., 2010; Mehu et al., 2012). In addition, whereas AUs associated with prototypical displays of sadness in both the upper and lower face (i.e., AU 1 + 4, AU 15) were ecologically valid cues of veracity, it was the muscles actions of the upper face (AU 1 + 4; inner brow raiser, combined with brow lowerer), but not the lower face (AU 15; lip corner depressor), which informed observers’ sympathy responses. Indeed, AU 1 + 4 may be a particularly useful signal for identifying the veracity of sadness as these muscles—innervated by the ipsilateral motor cortex—are under less voluntary motor control than the contralaterally innervated muscles of the lower face (Rinn, 1984). Furthermore, AU 1 + 4 may be a particularly salient cue because of its location on the face; people attend earlier and longer to the region around the eyes than the mouth when presented with a prototypical display of sadness (Calvo et al., 2018).

Highlighting a Methodological Advance

For decades, research has consistently documented poor discrimination accuracy of truths from lies in explicit lie detection paradigms (54%; Bond & DePaulo, 2006). Although most individuals are trusting of others (Levine, 2014), asking observers to make a truth-or-lie judgment immediately arouses suspicion and may impact the way they process and respond to others. Levine (2018) highlighted this issue as a difficult methodological hurdle to overcome in the pursuit of greater ecological validity. Here, we offer the study of emotional and behavioral reactions to liars (vs. truth tellers) as an alternative approach to studying responses to liars (vs. truth tellers). In other words, people may have distinct emotional and behavioral reactions to liars and truth tellers, even if unaware that the behavioral cues they are reacting to signal deception. In addition, these reactions can be assessed without arousing suspicion about deception among participants.

Limitations and Future Directions

Despite the knowledge gained from the current collection of studies, there are several potential avenues for future research. While we focus on emotional and behavioral responses to genuine versus deceptive expressers, cognitive operations are likely also at play, as predicted by the emotions as social information model (EASI; Van Kleef, 2009; Van Kleef & Côté, 2021). For example, although our participants were generally not suspicious about the presence of deception, research suggests that distrust can serve as a boundary condition for the effects of emotion on observer responses (e.g., Kim et al., 2017; Van Kleef et al., 2006). Future research should examine whether suspicion modulates the effect of expression veracity on observer reactions. Moreover, the focus of these studies was on only one emotional expression—that is, sadness—and how it elicits observer sympathy and intentions to donate. However, it is likely that a similar pattern of findings (e.g., Gunney & Ruben, 2015), indicating observer sensitivity to subtle cues to deception, would emerge for other emotions and their relevant social consequences. Indeed, there is already evidence that posed displays of happiness elicit less cooperation and posed anger is less effective in eliciting submission in negotiations (Côté et al., 2013; Grandey et al., 2005; Hennig-Thurau et al., 2006; Hideg & van Kleef, 2017; Tng & Au, 2014). Furthermore, while we have focused on cues to deception in the face, it is possible that observers also respond to emotional information inferred from other channels of communication, particularly verbal, vocal, or body language cues. Future research should examine whether emotional cues from other channels also influence observer behavior in predictable ways. Relatedly, the “behavior” studied here was hypothetical in nature. Although Kühberger et al. (2002) found that hypothetical and actual financial decisions closely approximate one another, future research would benefit from studying actual donations and other behavioral reactions to expresser sadness (e.g., comforting touch; Hertenstein et al., 2006). A further limitation of hypothetical donations studied here is the relatively large variance in observers’ donations across all studies. Observers in this study were not given a recommendation or “default” for a donation amount. The absence of benchmarks to guide their donations (e.g., see Azar, 2007) could explain the large variance in donations. Future research might consider including recommended donation amounts to anchor responses (e.g., Croson & Marks, 2001; Goswami & Urmsinsky, 2016), social norms (e.g., average donation amounts), or a smaller response range (e.g., US$0–US$10 vs. US$0–US$100) to reduce the range of responses.

Finally, it is important to note that studying observers’ emotional reactions to liars (vs. truth tellers) may be fruitful only in the context of emotional lies. Certainly, the social functional account of emotion provides predictions on observer emotional and behavioral reactions to various emotional expressions (Keltner & Haidt, 1999; Keltner & Kring, 1998), but provides less guidance on how emotionally neutral interactions are coordinated. As such, the confluence of theories presented here may provide little guidance for studying the effects of deception in less overtly emotional social interactions.

Conclusion

Bringing together findings on emotion and deception, we find that naive observers—in the absence of suspicion—experience less sympathy and offer less help to deceptive (vs. genuine) expressers of sadness and deceptive murderers (vs. genuine individuals) pleading for help to find a missing relative. Furthermore, observers used an empirically valid cue to deception to inform their emotional reactions: difficult-to-fake facial actions associated with sadness. Relying on
these emotional reactions to guide our social interactions may naturally protect observers from exploitation by emotional actors.

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) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs
Christopher A. Gunderson https://orcid.org/0000-0002-1218-2759
Leanne ten Brinke https://orcid.org/0000-0001-7080-9129

Supplemental Material
Supplemental material is available online with this article.

Notes
1. Additional sample demographics of Studies 1a, 1b, and 2 are included in the supplemental material.
2. Although we prefer setting power at .95, we opted for the conventional setting of .80 (Cohen, 1992) as .95 power required an extremely large sample size that was beyond our available resources.
3. See Supplemental Figure S1 for a replication of these effects.

References
Azar, O. H. (2007). The social norm of tipping: A review. Journal of Applied Social Psychology, 37(2), 380–402. https://doi.org/10.1111/j.0021-9029.2007.00165.x
Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. Psychological Science in the Public Interest, 20, 1–68. https://doi.org/10.1177/1529100619832930
Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. Psychological Inquiry, 2(2), 107–122. https://doi.org/10.1207/s15327965pi0202_1
Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. Personality and Social Psychology Review, 10, 214–234. https://doi.org/10.1207/s15327957pspr1003_2
Bond, C. F., Howard, A. R., Hutchison, J. L., & Masip, J. (2013). Overlooking the obvious: Incentives to lie. Basic and Applied Social Psychology, 35, 212–221. https://doi.org/10.1080/01973533.2013.764302
Brunswik, E. (1956). Perception and the representative design of psychological experiments (2nd ed.). University of California Press.
Calvo, M. G., Fernández-Martín, A., Gutiérrez-García, A., & Lundqvist, D. (2018). Selective eye fixations on diagnostic face regions of dynamic emotional expressions: KDEF-dyn database. Scientific Reports, 8, 1–10. https://doi.org/10.1038/s41598-018-35259-w
Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. Cognition and Emotion, 30, 1081–1106. https://doi.org/10.1080/02699931.2015.1049124
Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.
Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155–159. https://doi.org/10.1037/0033-2909.112.1.155
Côté, S., Hideg, I., & Van Kleef, G. A. (2013). The consequences of faking anger in negotiations. Journal of Experimental Social Psychology, 49, 453–463. https://doi.org/10.1016/j.jesp.2012.12.015
Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. Trends in Cognitive Sciences, 22, 388–399. https://doi.org/10.1016/j.tics.2018.02.006
Crivelli, C., & Fridlund, A. J. (2019). Inside-Out: From basic emotions theory to the behavioral ecology view. Journal of Nonverbal Behavior, 43, 161–194. https://doi.org/10.1007/s10919-019-00294-2
Croson, R., & Marks, M. (2001). The effect of recommended contributions in the voluntary provision of public goods. Economic Inquiry, 39(2), 238–249. https://doi.org/10.1111/j.1465-7295.2001.tb00063.x
DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. Psychological Bulletin, 111, 203–243. https://doi.org/10.1037/0033-2909.111.2.203
Dopson, W. G., Beckwith, B. E., Tucker, D. M., & Bullard-Bates, P. C. (1984). Asymmetry of facial expression in spontaneous emotion. Cortex, 20, 243–251. https://doi.org/10.1016/s0010-9452(84)80041-6
Dreyfuss, E. (2018, August 17). A bot panic hits Amazon’s Mechanical Turk. https://www.wired.com/story/amazon-mechanical-turk-bot-panic/
Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: A multimethod study. Journal of Personality and Social Psychology, 57, 55–66. https://doi.org/10.1037/0022-3514.57.1.55
Ekman, P. (2003). Darwin, deception, and facial expression. Annual New York Academy of Sciences, 1000, 205–221. https://doi.org/10.1196/annals.1280.010
Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). Duchenne smile: Emotional expression and brain physiology II. Journal of Personality and Social Psychology, 58, 342–353. https://doi.org/10.1037/0022-3514.58.2.342
Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiling. Journal of Nonverbal Behavior, 6, 238–252. https://doi.org/10.1007/bf00987191
Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. Motivation and Emotion, 10(2), 159–168. https://doi.org/10.1007/bf00992253
Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system: The manual on CD-ROM. Network Information
Ekman, P., Friesen, W. V., & O’Sullivan, M. (1988). Smiles when lying. Journal of Personality and Social Psychology, 54, 414–420. https://doi.org/10.1037/0022-3514.54.3.414
Ekman, P., Friesen, W. V., O’Sullivan, M., Chan, A., Diacoymanni-Thomas, I., Heider, K., Krause, R., LeCompte, W. A.,
Van Kleef, G. A., & Côté, S. (2021). The social effects of emotions. *Annual Review of Psychology, 73.* https://doi.org/10.1146/annurev-psych-020821-010855

Van Kleef, G. A., De Dreu, C. K., & Manstead, A. S. (2006). Supplication and appeasement in conflict and negotiation: The interpersonal effects of disappointment, worry, guilt, and regret. *Journal of Personality and Social Psychology, 91,* 124. https://doi.org/10.1037/0022-3514.91.1.124

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34,* 1–16. https://doi.org/10.1017/S0140525X10001354

Waller, B. M., Cray, J. J., Jr., & Burrows, A. M. (2008). Selection for universal facial emotion. *Emotion, 8,* 435–439. https://doi.org/10.1037/1528-3542.8.3.435

Wright Whelan, C., Wagstaff, G. F., & Wheatcroft, J. M. (2015). Subjective cues to deception/honesty in a high stakes situation: An exploratory approach. *The Journal of Psychology, 149,* 517–534. https://doi.org/10.1080/00223980.2014.911140

Zloteanu, M. (2019). *Emotion recognition and deception detection.* https://doi.org/10.31234/osf.io/crzne