ORIGINAL PAPER

Identifying the emotions behind apologies for severe transgressions

Samuel J. West[1,2] · Nicholas D. Thomson[2]

Abstract

Interpersonal transgressions, subsequent apologies, and offered (or withheld) forgiveness hold important consequences for both perpetrators and victims. Research has focused largely on the perceptions of victims and processes that promote forgiveness in relation to transgressions of low severity. In order to extend this domain of inquiry we examined the emotional substrates that facilitate and constrain apologies for severe transgressions (i.e., murder). We collected data on the final statements from incarcerated persons on death row and applied a sentiment analysis to obtain estimates of the emotions expressed in them (i.e., anger, anticipation, disgust, fear, joy, sadness, surprise, and trust). We manually coded each statement to indicate whether it exhibited some form of apology and compared the emotions expressed in apologetic statements versus non-apologetic statements. Results indicated that overall, final statements reflected high levels of joy and trust. Similar to studies of less severe transgressions, we found that 33.50% of statements contained some form of apology. Our comparative analyses revealed that apologetic statements reflected significantly greater sadness and less anger. In regression and subsequent dominance analyses we found that sadness and anger were the most important emotions to apologizing. We also found that anger moderated the association between sadness and the likelihood of apologizing such that as anger increased the effect of sadness decreased. Taken together these findings suggest that apologies for severe transgressions involve a delicate balance between sadness and anger rather than either emotion in isolation.

Keywords Apology · Interpersonal transgression · Violence · Sentiment analysis · Emotion

Introduction

Interpersonal conflict and the harms associated with it mark an unavoidable facet of the human social experience. Such harm can take many forms and occur at various levels of severity, but any such event carries distinct psychological consequences for both the perpetrator and target (e.g., Bastian et al., 2013). For example, aggressive provocateurs tend to derive feelings of hedonic pleasure during aggressive actions (e.g., Chester et al., 2019), whereas victims of peer aggression are more likely to experience feelings of depression (e.g., Söderberg & Björkqvist, 2020). Substantial research has examined the dispositional, physiological, and psychological precursors and antecedents of extremely harmful actions (e.g., murder). Despite this literature, we know little about factors that may elicit *apologies* from the purveyors of such severe transgressions.

The importance of apologies

Sincere apologies from transgressors serve as a powerful tool for ameliorating the negative psychological effects of inflicting harm on another person (e.g., Fisher & Exline, 2006). Such apologies are commonly marked by a clear acceptance of responsibility to the harm done on behalf of the transgressor and are devoid of defensive statements (Schumann, 2018; Schumann, 2014). Expressions of responsibility are considered a core component of apologies because they incur significant risk for the transgressor and return a sense of well-being to the aggrieved (Lawler, et al., 2003). Apologies also carry benefits for transgressors, as they increase victims' empathy for their assailants, make victims less likely to retaliate, and make victims more likely to forgive transgressions (Eaton, 2013; Ohbuchi et al.,

✉ Samuel J. West
   sawest@vsu.edu

1 Department of Psychology, Virginia State University,
   Petersburg, VA 23806, USA

2 Department of Surgery, Virginia Commonwealth University,
   Richmond, VA 23284, USA

1989). Despite these benefits, transgressors often refuse to apologize or offer otherwise insincere apologies. Seminal literature in this domain indicates that this phenomenon occurs primarily because transgressors deny the moral implications of their actions by rationalizing or refusing to take responsibility for them (e.g., Bandura, 2014; Baumeister et al., 1990).

The needs-based model of reconciliation argues that when one party harms another it creates distinct sets of emotional needs in the transgressor and the victim. Following a harmful action, transgressors suffer from feelings of diminished morality (Exline & Baumeister, 2000; Zechmeister & Romero, 2002) and guilt (Baumeister et al., 1994). The need-based model argues that as a result, transgressors develop a need for social acceptance which may be pursued by expressing related emotions to others to restore their moral identity which allows them to feel "rehumanized" (Shnabel & Nadler, 2008; Staub et al., 2005). The need for one's moral identity to be re-established can also be fulfilled by victims granting forgiveness. However, considerations of one's moral identity may also serve as a *barrier* to genuine apologies. Transgressors who do not see forgiveness as likely may not opt to apologize because proffering an apology without receiving forgiveness would serve to further impugn their moral identity (Schumann, 2018). This point is crucial to severe transgressions such as physical violence, as simple apologies are indeed less likely to be met with forgiveness, thus making the study of such apologies difficult. However, one specific context that may not be as impacted by such interpersonal dynamics are apologies offered at the end of one's life.

End of life apologies

Death carries with it an array of unique psychological experiences. One seminal model argues that people progress through five stages when approaching imminent death: denial, anger, bargaining, depression, and acceptance (Kübler-Ross, 1973). Individuals who are aware of their impending death (i.e., those in the 'acceptance' stage) may feel an urgent need to make amends for past wrongs before their life ends as a means of obtaining a 'good death' (Ko et al., 2015). As such, those who are clearly and explicitly aware of their upcoming death may be more motivated to re-establish their moral identities in their final moments as the only moral identity that will remain is in the memories of others. One specific population meets this criterion (i.e., is explicitly aware of their coming death) and has a record of severe transgressions: death row inmates. Death row inmates are intimately aware of their pending fates. People who are incarcerated and have been condemned to die are trapped in a constant rumination about their own demise, an experience

described as 'living death' (Johnson, 1979). As a result, some may be highly motivated to apologize for the actions they were convicted of in order to re-establish their sense of a moral identity in their final moments. However, in competition with this possible motive are various antagonistic traits that are elevated in such populations and negatively linked with accepting responsibility or showing remorse (Goldstein et al., 2006; Spice et al., 2015; West & Chester, 2021). Such individuals are at a greater risk of engaging in antisocial behavior and thus severe transgressions, such as violence.

Emotions, apologies, and antisocial behavior

Much of the extant literature indicates that disturbances in the recognition and experience of the full breadth of human emotion underlies antisocial tendencies, particularly in relation to harming others (e.g., Marsh, 2013; Thomson et al., 2021). One theory argues that antisocial behaviors develop over time due to a deficit in the ability to experience or recognize fear (e.g., Newman et al., 2010; Sylvers et al., 2011). Other research places a major emphasis on differences in anger, indicating that those with antisocial dispositions are more likely to exhibit explosive outbursts of anger, or positive urgency, indicating that positive affective experiences drive some to behave more impulsively (Gray et al., 2019; Hicks & Patrick, 2006). Nevertheless, such individuals do at times apologize. This may be particularly likely for so-called "successful psychopaths" who have improved abilities of social cognition and impulse control which may allow them to discern when apologies may be beneficial personally (Cangemi & Pfohl, 2009; Lasko & Chester, 2021; Lasko et al., 2019). Despite the body of literature examining the involvement of emotionality in antisocial behavior, no known work has examined the role of such emotions in promoting the expression of apology among severe transgressors. These aspects of whether one apologizes are likely to be qualified by the emotional state of the transgressor.

The broader literature on emotion and behavioral motivation indicates that each discrete emotion is linked with a combination of arousal (i.e., high vs. low) and affective valence (i.e., positive vs. negative; Christie & Friedman, 2004). As such, certain emotions (e.g., anger, happiness) are approach oriented—they make a person more likely to engage in a given behavior (e.g., Harmon-Jones & Allen, 1998). In contrast, some emotions are withdrawal oriented (e.g., disgust, fear) and evoke automatic physical withdrawal from a given stimuli or action (e.g., Pond et al., 2012). Given that emotional experiences serve as a primary motivating factor behind most behavior, it would seem a key starting point in understanding any emotion-laden behavior such as delivering an apology (LeDoux, 2012). Consistent with this

view, recent work indicates that state empathy (rather than trait empathy) predicts more comprehensive and genuine apologies for interpersonal transgressions (Schumann & Dragotta, 2021). However, the literature examining the precedents of apologies has yet to explore the specific involvement of individual emotions as experienced in-the-moment by transgressors. Considering the evidence that genuine apologies from transgressors carry significant benefits for all parties, understanding the emotions that motivate such actions may be an important key to facilitating genuine apologies, thereby fostering an improved approach to reconciliation (Eaton, 2013; Fisher & Exline, 2006; Ohbuchi et al., 1989). Research examining the final statements of death row inmates and the blogs of terminally ill individuals reveals that these experiences are positive in affective valence (Goranson et al., 2017). Extending this work by examining the specific emotions expressed by those facing execution may further reveal insights regarding the psychological and human costs of the death penalty.

Current study

The facilitating factors and barriers to apologies are relatively well known for transgressions ranging low to moderate in severity. However, little research has examined the emotional experiences involved in apologies following *severe* transgressions. In order to address this gap in the literature the current study extracted the recorded final statements of incarcerated persons who were executed by the state of Texas and subjected each statement to a sentiment analysis to explore the emotional profiles exhibited in the final statements of those convicted of murder or conspiracy to murder. We then compared the profiles of those who apologized during their final moments against those who did not to understand the emotional experiences that are most important for apologizing over severe transgressions. We did not have any specific hypotheses for the current study as it was exploratory in nature and our analyses were not pre-registered.

Methods

Data acquisition

All data were acquired from the Texas Department of Criminal Justice (TDCJ) website which hosts publicly available profiles of all executions from December 1982 to the present. In total, this repository contained information about 573 executions conducted by the state of Texas (as of 3/16/2022). Capital crimes in the state of Texas all involve murder or conspiracy to murder (i.e., hiring another person

to kill). Individual case profiles included information about participant's race, age, education, and their last statements. All executions occurred in the same facility: the so-called Huntsville Unit.

Data cleaning

Of the initial set of 573, 139 cases were excluded because their records contained no final statement, or their recorded final statement was simply them declining to make a statement. We then examined each statement manually to ensure text that was not relevant to each statement was removed. This was necessary because the individuals tasked with recording the statements sometimes added narrative text describing what was happening (e.g., "he mouthed 'love you' to his mother") or made an explicit note that text had been omitted due to profanity or was unintelligible. All such text was removed prior to our sentiment analysis. In some cases, final statements were written and distributed to those in attendance. In these cases, we retained the text of these written final statements. Our final dataset contained 435 statements. However, eight of these cases were not included in our analyses due to the discovery of exonerating evidence post-execution per the Death Penalty Information Center (https://deathpenaltyinfo.org/policy-issues/innocence/executed-but-possibly-innocent). We excluded these cases as the emotions and final statements of innocent persons are likely linked to distinct psychological experiences in relation to the rest of our sample, resulting in a final sample of 427.

Statistical power statement

The sample used in the current work was not determined using any *a-priori* procedure as we included all usable cases from the TDCJ database. A sensitivity analysis using G*Power version 3.1.9.7 indicated our sample had 80% power to detect effects larger than $r = 0.13$ (Faul et al., 2009).

Coding procedure

The first author and a trained research assistant independently read and coded each statement for whether it contained an apology, similar to prior research utilizing earlier versions of the TDCJ data (e.g., Eaton & Theuer, 2009). Apologies were defined as any statement that reflected an acceptance of personal responsibility for the transgression (e.g., "…I am sorry for the pain and suffering I have caused you…"). Some statements contained appeals for forgiveness from the victim's family, but did not contain apologies and were thus not coded as apology statements. We did not include these statements as apologies because such statements did not involve the core components of apology (e.g., accepting responsibility for doing harm) but still asked

for forgiveness from the aggrieved. Although such appeals may sometimes result in forgiveness, the core components of sincere apologies identified in the literature are absent (Schumann, 2018).

Data analysis

All analyses were conducted using R version 4.1.1. We examined the interrater reliability of our statement coding using the *irr* package (Gamer et al., 2019). All sentiment analyses were conducted using the sentimentR package (Rinker, 2021). We also estimated a logistic regression model using the base R *glm* function and conducted dominance analyses using the dominanceAnalysis package (Navarrete & Soares, 2020) to examine the emotional experiences that were most important to apologies for severe transgressions. Finally, we probed the interaction term from our exploratory moderation model using the *simple_slopes* function from the reghelper package (Hughes & Beiner, 2021). Given the exploratory nature of this work we applied the Benjamini–Hochberg correction to all *p*-values presented in this manuscript (and supplemental analyses; 84 values in total) in order to appropriately control the false discovery rate (FDR; Benjamini et al., 2009).

Sentiment analysis

Sentiment analysis allows the researcher to extract quantitative estimates of the affect and emotions expressed in spoken or written language. Many approaches to sentiment analysis exist, but recent years have seen the introduction of open-source automated sentiment analysis software packages (see Naldi, 2019 for a review of R packages for sentiment analysis). The majority of such packages (and our application specifically) utilize a "bag-of-words," lexicon-based approach. That is, individual words are compared against an extensive lexicon that contains values for affective valence and emotional content. These values are then combined to produce an overall score for the variables requested (i.e., positive/negative affect or specific emotions). However, one common issue that affects many sentiment analysis packages is that they do not account for sentence modifiers or negators. For example, many such packages would assign a similar score to the sentences "I am so happy" and "I am so not happy". We chose the sentimentR package for our sentiment analysis because it is the only existing R package that accounts for negators and modifiers by default. sentimentR accomplishes this by examining the six words surrounding each individual term before assigning an ultimate score (Rinker, 2021). This approach allowed us to yield estimates of the levels of specific emotions expressed in each statement. The sentimentR package implements the psychoevolutionary theory of emotion which posits eight core emotions to the human experience: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust (e.g., Plutchick, 1980). As such, we utilized estimates of each of the emotions extracted by the sentimentR package using the NRC Emotion dictionary (Mohammad & Turney, 2013). These values were estimated such that they were standardized to statement length, allowing us to contrast the emotions of shorter statements with longer statements. Recent research indicates that such automated sentiment analysis approaches are roughly equivalent to the accuracy of trained human coders (Provoost et al., 2019).

Data availability

All data and code needed to reproduce our findings are publicly available on the open science framework (https://osf.io/ad72u/files/).

Results

All sample demographic information (i.e., race, education, sex, occupation, and age) is presented in Table 1. Descriptive statistics of the emotion estimates produced by our sentiment analysis are presented in Table 2.

Note. Only 139 cases had data for sex in the TDJC database.

Each of the study variables had some univariate outliers (i.e., ± 3 SD) which were Winsorized. Final statements

Table 1 Sample demographic information

Race, % (n)	
Black	36.80% (157)
Other	0.50% (2)
Hispanic	20.60% (88)
White	42.20% (180)
Age	
M (SD)	39.37 (8.57)
Range	24—70
Years of Education	
M (SD)	10.11 (2.03)
Range	3—16
Occupation Type	
Laborers	84.60%
Law Enforcement	1.00%
Management	1.60%
Office Work	3.90%
Unknown	8.90%
Sex, % (n)	
Male	98.56% (137)
Female	1.44% (2)

Table 2 Emotion estimate descriptives

Variable	*M*	*SD*	Range	Outliers
Anger	0.02	0.03	0.00–0.33	8
Anticipation	0.04	0.05	0.00–0.67	5
Disgust	0.01	0.02	0.00–0.22	6
Fear	0.03	0.03	0.00–0.33	8
Joy	0.06	0.05	0.00–0.33	8
Sadness	0.02	0.02	0.00–0.12	7
Surprise	0.01	0.02	0.00–0.17	6
Trust	0.05	0.05	0.00–0.33	7

were 108.98 words long on average (*SD* = 104.48). Zero-order bivariate correlations among the sentiment estimates are presented in Table 3. Example statements characteristic of those high in a given emotion (e.g., anger) are available in Supplemental Document 2 (Table S8).

Interrater reliability

We tested the interrater reliability of our apology coding by examining the interclass correlation coefficient (ICC). We used a one-way model for single measurement units testing the consistency of our ratings. This analysis revealed substantial consistency in coding, *ICC(1)* = 0.93, 95% CI 0.92, 0.94, indicating strong reliability between coders (Koo & Li, 2016). The original coding of the first author was thus retained. In respect to apologies, 143 (33.50%) of those in the current sample apologized for their actions whereas the remaining 284 (66.50%) did not.

Emotional profiles of final statements

We applied a series of paired-samples *t*-tests to explore the general emotional profile expressed in all statements. All results of these analyses are detailed in Supplemental Document 1 (Tables S1–S7). In general, final statements delivered prior to execution reflected particularly high levels of joy and trust (Fig. 1), as these emotions were significantly

higher than all others. The emotions expressed the least were disgust, surprise, anger, and sadness. In terms of frequency, statements expressed 5.76 emotions on average (i.e., the number of emotions estimated as above zero).

To further understand the emotional experiences underlying apologies, we conducted a series of independent *t*-tests comparing the mean level of each emotion expressed between those who apologized and those who did not. These analyses (Table 4) revealed that those who included apologies in their final statements expressed significantly less anger than those who did not (Fig. 2). Those who apologized also expressed less joy and more sadness than those who did not. Finally, those who apologized expressed greater surprise and less trust than those who did not. However, fear, disgust, and anticipation did not differ between these groups.

Identifying the most important emotions behind apologies

To identify which of these emotions was most strongly linked with apology, we applied a bootstrapped dominance analysis procedure to a binomial logistic regression model. First, we modeled the dichotomous apology variable as the outcome and each of the eight emotions as independent variables in a multiple logistic regression. This model correctly classified 69% of statements and indicated that expressions of anger and joy were both linked with a lower likelihood of apologizing (Nagelkerke R^2 = 0.20; AICc = 496.05). In contrast, sadness and surprise were the only emotions significantly linked with a greater likelihood of apologizing. Based on these findings, it appeared that sadness was the strongest positive predictor of apologizing for a severe transgression, where each unit increase in sadness was associated with an 85% increase in the likelihood of apologizing (Table 5). In contrast, anger appeared to be the strongest *negative* predictor of apologizing (Fig. 3), as each unit increase in anger was linked with an approximate 76% reduction in likelihood of apologizing.

In order to confirm which of these variables was the most important predictor we subjected this model to a dominance

Table 3 Zero-order correlations among all emotion variables

	Variable	1	2	3	4	5	6	7
1	Anger	–						
2	Anticipation	0.08	–					
3	Disgust	0.40***	0.09	–				
4	Fear	0.65***	0.38***	0.37***	–			
5	Joy	− 0.24***	0.43***	− 0.08	0.08	–		
6	Sadness	0.45***	0.04	0.41***	0.50***	− 0.24***	–	
7	Surprise	0.13*	0.44***	0.27***	0.29***	0.25***	0.30***	–
8	Trust	0.10	0.62***	0.23***	0.38***	0.50***	− 0.00	0.35***

*p < 0.05, **p < 0.01, ***p < 0.001

Fig. 1 Average levels of each emotion expressed in last statements. Error bars reflect 95% confidence intervals
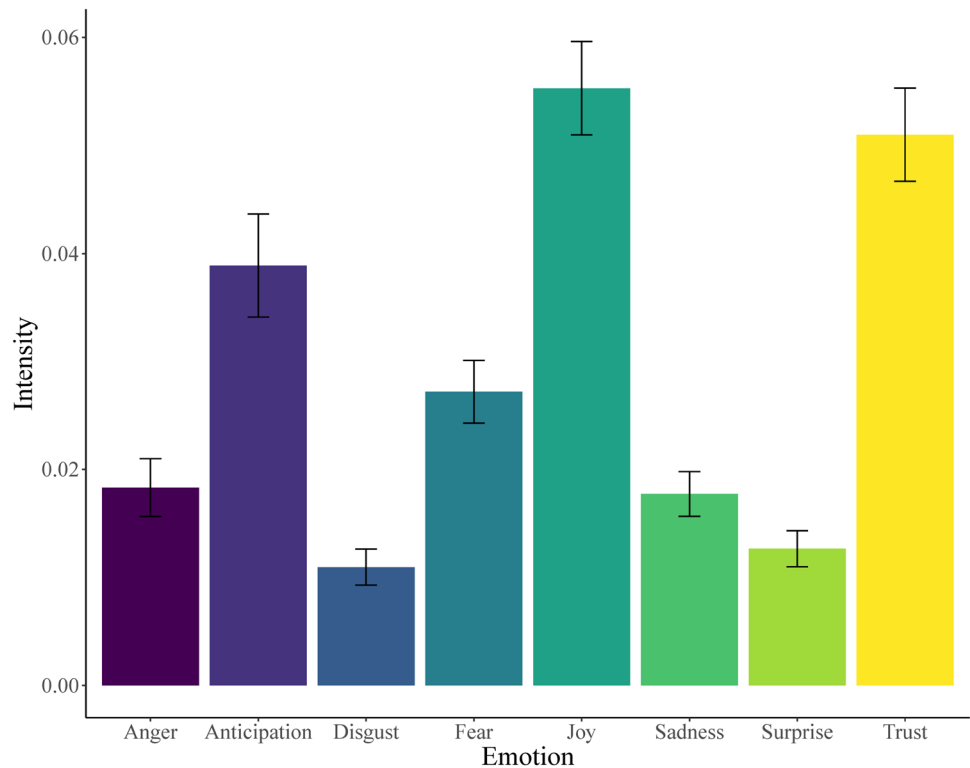

Table 4 Independent-samples t-tests Comparing Levels of Emotions Expressed between Those Who Apologized and Those Who Did Not

Emotion	t	p	d	95% CI
Anger*	− 2.55	0.015	− 0.23	− 0.41, − 0.05
Anticipation*	− 1.35	0.198	− 0.12	− 0.30, 0.05
Disgust	0.68	0.514	0.07	− 0.13, 0.27
Fear*	0.01	0.994	0.00	− 0.18, 0.18
Joy*	− 3.52	<0.001	− 0.33	− 0.51, − 0.14
Sadness	4.21	<0.001	0.43	0.23, 0.63
Surprise	1.93	0.067	0.20	− 0.00, 0.40
Trust*	− 2.02	0.056	− 0.19	− 0.37, 0.01

*Groups differed in variability as indicated by a significant Levene's test, Welch's t and corresponding effect size reported

analysis using a 5,000-sample non-parametric bootstrap procedure. Full results of our dominance analysis and subsequent bootstraps are presented in Supplemental Document 2 (Table S9). This analysis revealed that sadness was the single most important emotion in predicting whether final statements contained apologies, as it exhibited complete dominance as a predictor above and beyond all other emotions except for anger, where conditional dominance was established in favor of sadness. We further examined the conditional dominance of our predictors by plotting their contribution values against the level of model complexity (Fig. 4). Sadness emerged as the dominant predictor of apologies

excepting at the highest levels of model complexity where anger emerged as the stronger predictor. That is, when all other emotions were accounted for, anger accounted for the most variance in apologies and an individual predictor.

Exploring the roles of sadness and anger

Given the importance of anger and sadness to apology, we compared the different levels of these emotions within each statement type. A paired samples t-test among those who apologized revealed that their statements expressed significantly more sadness than anger, $t(142)=5.88$, $p<0.001$, $d=0.49$, 95% CI = 0.33, 0.67. Despite this difference, a one-sample t-test revealed that apologetic statements still contained levels of anger significantly greater than zero, $t(142)=11.82$, $p<0.001$, $d=0.98$, 95% CI 0.79, 1.18, as 71.30% of apologetic statements expressed some degree of anger. In contrast, those who did not apologize demonstrated significantly greater levels of anger than sadness, $t(283)=3.00$, $p=0.004$, $d=0.18$, 95% CI 0.07, 0.30. Considering these results and our zero-order correlations indicating the positive association between anger and sadness, we conducted an exploratory moderation analysis to test if anger and sadness interacted to predict apology. This model accounted for approximately 17% of the variability in apologies, AICc=497.55, Nagelkerke's $R^2=0.17$. Like our initial regression model, we found sadness exhibited a significant positive effect, OR 2.66, 95% CI 1.96, 3.71, $p<0.001$, and anger exhibited a negative effect, OR 0.49, 95%

Fig. 2 Mean levels of emotions split by apology group. Error bars represent 95% confidence intervals
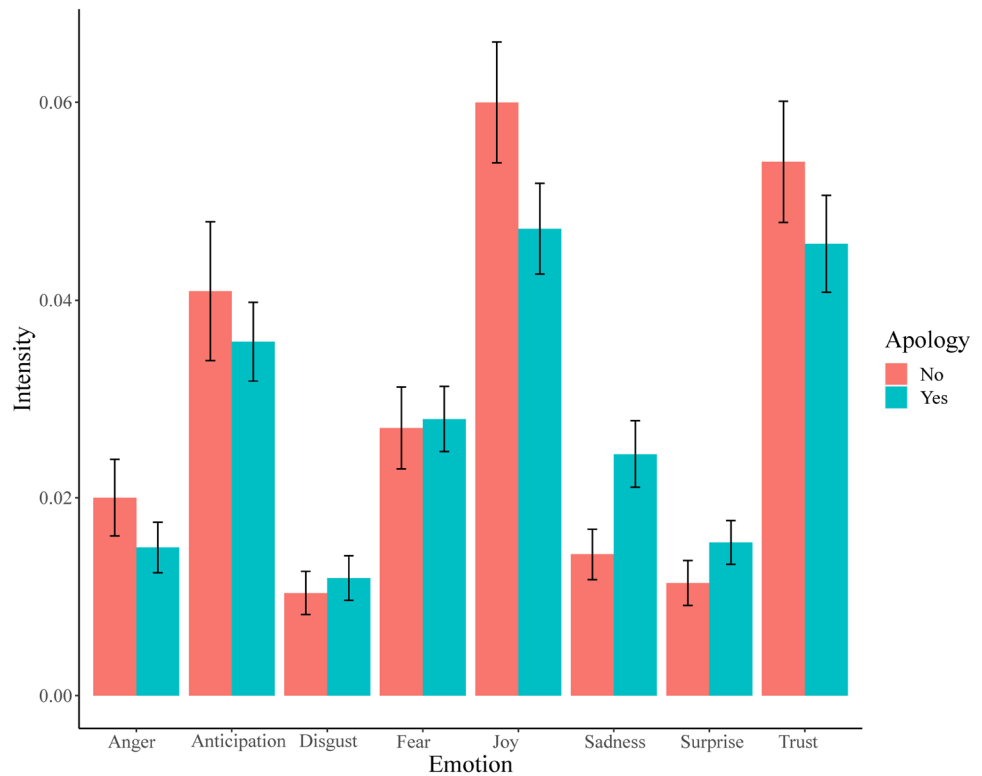

Table 5 Binomial logistic regression results

Emotion	OR	95% CI	p
Anger	0.24	0.13 – 0.40	< 0.001
Anticipation	0.84	0.50 – 1.36	0.509
Disgust	1.22	0.86 – 1.72	0.296
Fear	1.52	0.94 – 2.49	0.108
Joy	0.58	0.39 – 0.83	0.006
Sadness	1.85	1.33 – 2.65	< 0.001
Surprise	1.62	1.16 – 2.29	0.007
Trust	0.83	0.55 – 1.22	0.382

OR Odds Ratio

CI 0.30, 0.73, $p = 0.002$, on apology. However, these effects were qualified by their interaction, *OR* 0.65, 95% CI 0.47, 0.87, $p = 0.009$. Decomposition of this interaction (Table 6) revealed that the strength of the relationship between sadness and likelihood of apology decreased as levels of anger increased.

Discussion

Research provides significant insight into the psychological processes involved in interpersonal transgressions and subsequent apologies. However, such research largely focuses on lesser interpersonal transgressions (e.g., lying), with little known work focusing specifically on severe transgressions (e.g., murder). We provided a preliminary exploration of the emotional experiences that facilitate and constrain apologizing for such actions. The current work revealed that two emotions specifically, sadness and anger, played distinct roles in motivating apologies.

Emotional profiles of final statements

Our sentiment analyses and *t*-tests revealed that final statements were generally marked by high expressions of joy and trust, moderate levels of fear and anticipation, and low levels of surprise, anger, disgust, and sadness. These findings are broadly consistent with prior research that has applied similar techniques to the final statements from the TDCJ's database. Specifically, Goranson et al. (2017) applied a linguistic analysis to the statements contained in the TDCJ database which included executions that occurred 1982–2013. They found that the affective content of final statements was significantly more positive than negative (Goranson et al., 2017; Study 2). Our results replicated and extended these findings by examining the specific emotions that underlie the affective profile observed by previous research and linking these emotions to explicit apologies. Goranson et al. also found that the blog posts of terminally ill patients exhibited an extremely similar affective profile to the final statements (2017; Study 1). Finally, our finding that 33.50% of final statements contained apologies is consistent with research indicating that romantic partners typically engage in sincere
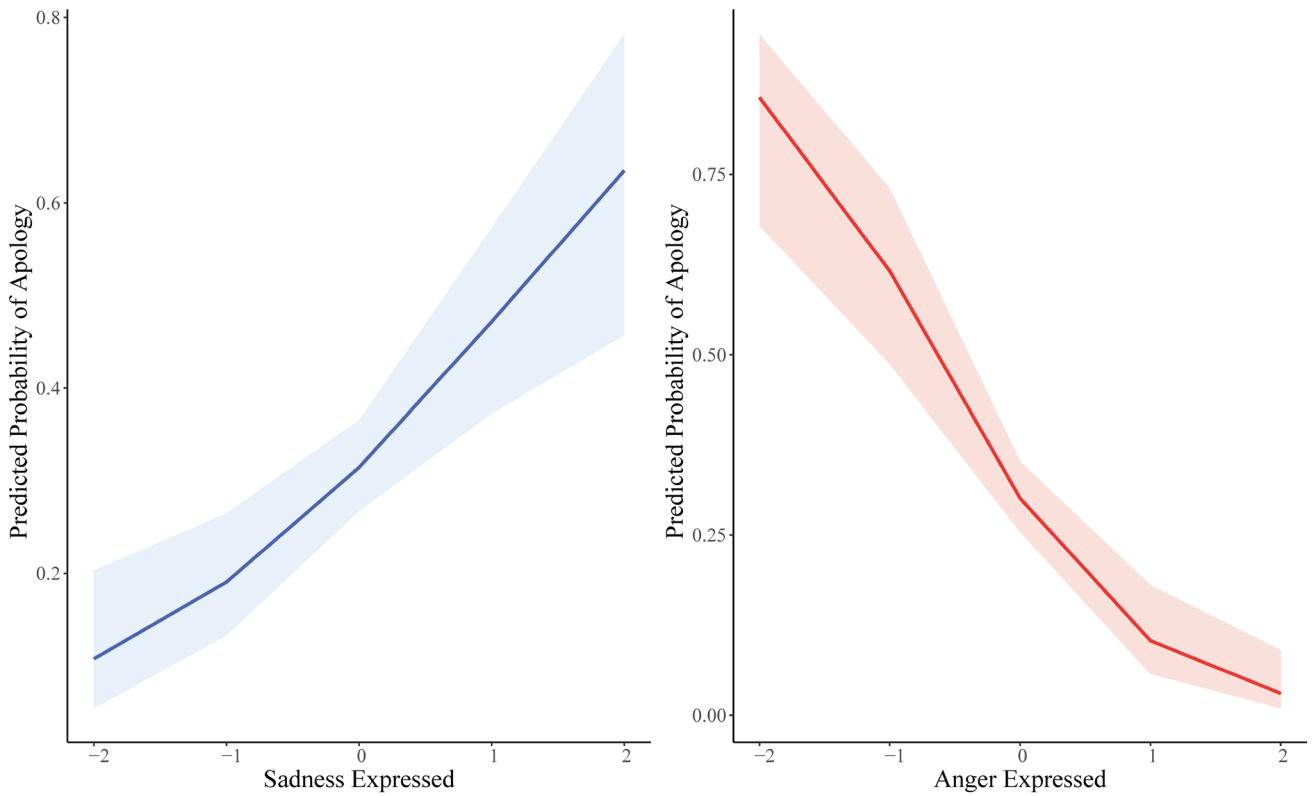
Fig. 3 Predicted probability of apologizing as a function of expressed sadness (left) and anger (right). X-axes reflect standardized values

Fig. 4 Proportion of variance explained by each predictor (via Nagelkerke's R^2) emotion displayed as a function of the level of model complexity. Model complexity is indexed by the number of predictors (i.e., emotions) included in the subset models summarized here
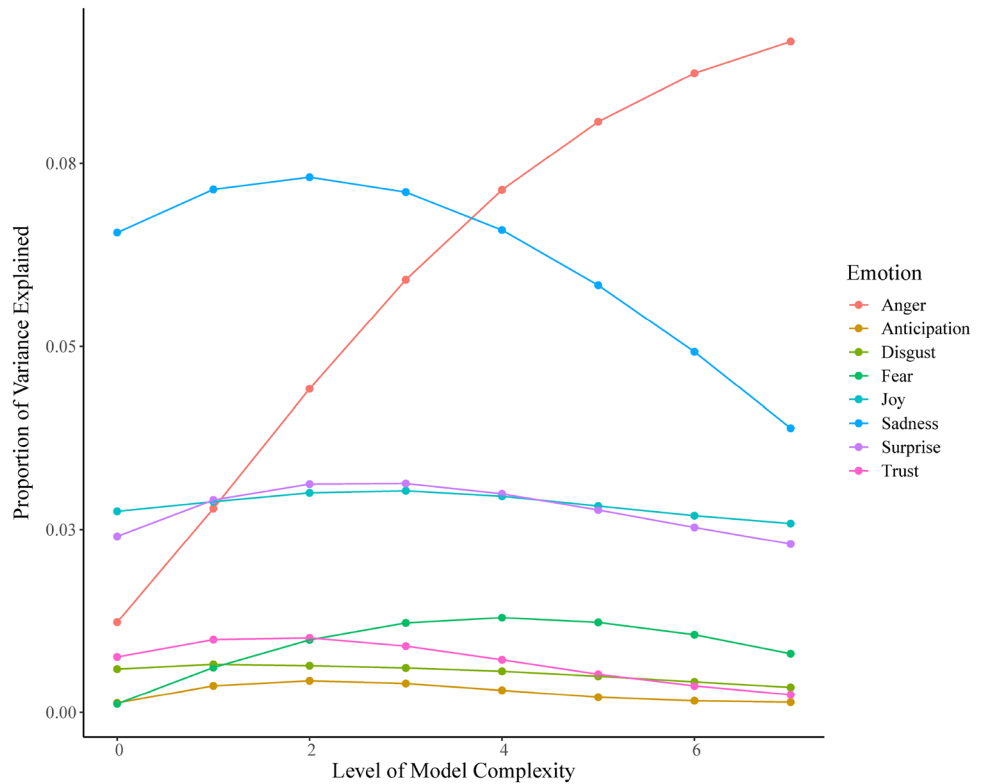
Table 6 Conditional effects of sadness and anger on apology

Level of emotion	Slope estimate	SE	t	p
Anger				
Low	1.31	0.22	5.68	< 0.001
Moderate	1.00	0.16	6.10	< 0.001
High	0.68	0.18	3.88	< 0.001
Sadness				
Low	− 0.31	0.29	− 1.09	0.302
Moderate	− 0.71	0.22	− 3.23	0.002
High	− 1.12	0.24	− 4.62	< 0.001

apologies are a similar rate (i.e., 30.94% of the time; Schumann, 2012). As such, it seems likely that the current findings are not unique to the psychological experiences of people incarcerated for murder.

Emotions and apologies

Our independent-samples *t*-test results indicated that distinct profiles of emotion existed between those who apologized and those who did not. Specifically, we observed that apologies were typified by elevated levels of sadness and surprise, but lower levels of anger, joy, and trust. These findings are broadly consistent with work indicating that emotional states are of critical importance when it comes to apologies (Schumann & Dragotta, 2021). Critically, our logistic regression and subsequent dominance analysis indicated that sadness was the single most important positive predictor of apology. This finding is consistent with research indicating that expressions of sadness in written and verbal admissions are generally an indicator of genuine remorse (Villar et al., 2014). In contrast, we found anger was the most important negative predictor of apology. Anger's negative impact on the likelihood of apology is consistent with prior work indicating other-blame as a crucial feature of anger (e.g., Ben-Zur & Breznitz, 1991; Levine, 1996; Vansteelandt & Van Mechelen, 2006). Our findings are also consistent with recent research indicating that anger alone predicts selfish outcomes in social moral dilemmas, but that similar levels of sadness and anger are more closely linked to selfless outcomes (Lutz & Krahé, 2018; Plaks et al., 2021). These findings considered jointly yield an important insight into the emotional substrates of apologizing. Although sadness and anger are both negative in valence, they differ in level of arousal and directional motivation, as anger marks a high-arousal, approach-oriented emotional state, where sadness is a low-arousal state associated with behavioral withdrawal (Christie & Friedman, 2004; Feldman, 1995; Harmon-Jones & Allen, 1998). As such, it seems plausible that some degree of both emotions may facilitate apology.

Our analyses indicated that apologetic statements contained greater expressions of sadness than anger, but that such statements still contained substantial anger expression. Similarly, our zero-order correlations revealed that sadness and anger were positively correlated. These results are perhaps best understood through the lens of a socio-functional account of emotion (e.g., Keltner & Gross, 1999). In this context, sadness may cue transgressors to an opportunity to make amends and thus drive them to seek forgiveness. However, sadness alone is generally associated with a lack of behavioral activation and thus some degree of anger may contribute arousal to one's internal state thus facilitating an ultimate apology (e.g., Christie & Friedman, 2004). In other words, sadness appears to be a critical component of motivating apologies for severe transgressions because of its affective valence, and anger appears to be a critical component due to its contribution of arousal. Results from our moderation analysis suggested that a careful balance between sadness and anger is important, as the effect of sadness decreased as anger levels increased. These findings also hold implications for clinical approaches to addressing antisocial behavior.

Clinical implications

Substantial research indicates that those inclined towards antisocial behaviors commonly exhibit low levels of psychological and physiological arousal even in the face of threatening stimuli (e.g., Thomson et al., 2021). Such individuals commonly engage in sensation seeking behaviors and are typified by a propensity to boredom and explosive bouts of anger (Gray et al., 2019; Hicks & Patrick, 2006; Pfattheicher et al., 2021). As such, it may be that those with antisocial tendencies experience sadness as a more aversive emotional state than anger because of the difference in arousal. As a result, the functionality of sadness may become disrupted because the aversiveness of the low arousal state is misattributed to this feeling and the act of apologizing itself. When confronted by negative emotional experiences the individual may thus have a strong implicit preference towards anger as opposed to chronic under-arousal. This interpretation is consistent with research indicating that people do preferentially pursue negative emotional experiences for utilitarian purposes (e.g., Riediger et al., 2009) and that affective experiences are commonly attributed to unrelated stimuli (e.g., Payne et al., 2005). However, experimental work indicates that induced sadness can reduce the effects of anger on antisocial behaviors (i.e., aggression; Lutz & Krahé, 2018; Zhan et al., 2015). As such, clinical researchers may find the development of such experimental inductions into clinical interventions for those prone to antisocial behaviors that place an emphasis on low-arousal tolerance a fruitful area for future research.

The psychological and human costs of the death penalty

Although the primary interest of this investigation were the emotions underlying apologies for severe transgressions, the nature of our data provided some insight into the societal costs and benefits of the death penalty. The death penalty is generally defended as a state-mandated means by which victims of severe transgressions and their families can obtain closure (Zimring, 2004). Such arguments hinge on the assumption that executing a person will provide a psychological benefit to the victim's family and will act as a deterrent for any would-be offenders. However, research has found that capital punishment does not result in deterrence for homicide or other forms of antisocial behavior (Hong & Kleck, 2018; Peterson & Bailey, 1991), and that the execution of death row inmates rarely has psychological benefits for the witnesses and other co-victims (Eaton & Christensen, 2014; Goodwin, 1997). In contrast, findings from the current study and prior research indicate that the emotions expressed most strongly among final statements was joy or positive affect more generally (Goranson et al., 2017). Given that a similar effect has been found among blog posts of terminally ill patients (Goranson et al., 2017) it would seem that execution marks a final escape from the extended suffering faced by those on death row (e.g., Johnson, 1979). Despite such suffering, there is little evidence to indicate that these features of the legal system are helpful to victims' families. Evidence indicates that mediated meetings between victims and transgressors result in a similar rate of transgressor apologies as those found in our study (i.e., 38%), but that such discussions also improved the psychological wellbeing of 58% of participating victims (Umbreit et al., 2003). A meta-analysis examining the impact of such mediated reconciliation between victims and transgressors indicates that participation in such reconciliation is linked with a significantly reduced rate of future transgressions (Nugent et al., 2003). Taken together, our findings and prior research suggest that a thorough investigation of whether the psychological and human costs of the death penalty outweigh the ostensible societal benefits is warranted.

Limitations

The current study provided an in-depth look at the emotions that drive apologies for the most severe of interpersonal transgressions: murder. Despite the novel nature of this research, our findings must be considered in light of several limitations. First and foremost, the raw text of last statements we obtained from the TDCJ database was transcribed by a third party and we thus must assume the accuracy of these transcriptions. Similarly, we did observe evidence of non-standard transcription methods in the raw dataset as some entries included third-person narrative text (e.g., "He said okay") instead of simply the words spoken. Although we did clean each statement individually to remove such text, its presence suggests that other features could have differed across transcribers that could have been more difficult to spot. Complicating matters further, there were several cases where it appeared that the speaker was executed before they finished delivering their final statements, thus making it impossible to know what the remainder of their statement(s) would have contained. Second, the statements collected here were delivered under extreme circumstances that involved many persons. Relatedly, no data were available to indicate who attended the execution as an observer, making it impossible to determine who each statement was directed towards (i.e., the victim's family or their own family). Such circumstances very likely impacted the behavior of speakers, but the degree of consistency of our findings with other studies of apology suggest that this is not likely the case in terms of our inferences about the involvement of emotions. Third, our analyses are exploratory in nature and rely on retrospective data and correlational analyses. As such, we are unable to make any inferences regarding the causal role any such emotions may play in apologies for severe transgressions. Future research should seek to replicate our findings in samples of non-incarcerated persons to confirm and further explore the roles of anger and sadness in offering apologies for interpersonal transgressions in a controlled laboratory environment.

Conclusion

Interpersonal conflicts and transgressions are a fact of social life. Such harmful actions hold distinct psychological consequences for transgressors and victims alike, whereas the acceptance of genuine apologies hold positive benefits for both parties. In this study we explored the emotional profiles of final statements from inmates facing execution and compared the profiles of those who apologized for their actions against those who did not. We found that sadness and anger were the two most important emotions to apology, but that they interacted with each other. Our findings thus suggest that human emotions may facilitate apology by functionally recruiting the arousal components of anger and affective valence of sadness to provide the behavioral activation necessary to apologize to a group of people. We also found that those in our sample apologized for their severe transgressions at similar levels as apologies for less severe transgressions. Taken together, our findings could indicate that those prone to antisocial behaviors don't necessarily lack the ability to experience remorse, but rather are highly motivated to avoid the emotional state(s) that accompany it.

References

Bandura, A. (2014). Moral disengagement in the perpetration of inhumanities. In A. Bandura (Ed.), *Perspectives on evil and violence* (pp. 193–209). Psychology Press.

Bastian, B., Denson, T. F., & Haslam, N. (2013). The roles of dehumanization and moral outrage in retributive justice. *PLoS ONE, 8*(4), e61842. https://doi.org/10.1371/journal.pone.0061842

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*(2), 243–267. https://doi.org/10.1037/0033-2909.115.2.243

Baumeister, R. F., Stillwell, A., & Wotman, S. R. (1990). Victim and perpetrator accounts of interpersonal conflict: Autobiographical narratives about anger. *Journal of Personality and Social Psychology, 59*(5), 994–1005. https://doi.org/10.1037/0022-3514.59.5.994

Benjamini, Y., Heller, R., & Yekutieli, D. (2009). Selective inference in complex research. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences, 367*(1906), 4255–4271. https://doi.org/10.1098/rsta.2009.0127

Ben-Zur, H., & Breznitz, S. (1991). What makes people angry: Dimensions of anger-evoking events. *Journal of Research in Personality, 25*(1), 1–22. https://doi.org/10.1016/0092-6566(91)90002-8

Cangemi, J. P., & Pfohl, W. (2009). Sociopaths in high places. *Organization Development Journal, 27*(2), 85–96.

Chester, D. S., DeWall, C. N., & Enjaian, B. (2019). Sadism and aggressive behavior: Inflicting pain to feel pleasure. *Personality and Social Psychology Bulletin, 45*(8), 1252–1268. https://doi.org/10.1177/0146167218816327

Christie, I. C., & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: A multivariate approach. *International Journal of Psychophysiology, 51*(2), 143–153. https://doi.org/10.1016/j.ijpsycho.2003.08.002

Eaton, J. (2013). The effects of third-party validation and minimization on judgments of the transgressor and the third party. *British Journal of Social Psychology, 52*(2), 273–289. https://doi.org/10.1111/j.2044-8309.2011.02066.x

Eaton, J., & Theuer, A. (2009). Apology and remorse in the last statements of death row prisoners. *Justice Quarterly, 26*(2), 327–347. https://doi.org/10.1080/07418820802245078

Exline, J. J., & Baumeister, R. F. (2000). Expressing forgiveness and repentance. *Forgiveness: Theory Research, and Practice, 133155*, 133–155.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Feldman, L. A. (1995). Variations in the circumplex structure of mood. *Personality and Social Psychology Bulletin, 21*(8), 806–817. https://doi.org/10.1177/0146167295218003

Fisher, M. L., & Exline, J. J. (2006). Self-forgiveness versus excusing: The roles of remorse, effort, and acceptance of responsibility. *Self and Identity, 5*(02), 127–146. https://doi.org/10.1080/15298860600586123

Gamer, M., Lemon, J., & Singh, I. (2019). irr: Various coefficients of interrater reliability and agreement. R package version 0.84.1, https://cran.r-project.org/web/packages/irr/

Goldstein, R. B., Grant, B. F., Huang, B., Smith, S. M., Stinson, F. S., Dawson, D. A., & Chou, S. P. (2006). Lack of remorse in antisocial personality disorder: Sociodemographic correlates, symptomatic presentation, and comorbidity with Axis I and Axis II disorders in the National Epidemiologic Survey on Alcohol and Related Conditions. *Comprehensive Psychiatry, 47*(4), 289–297. https://doi.org/10.1016/j.comppsych.2005.11.002

Goranson, A., Ritter, R. S., Waytz, A., Norton, M. I., & Gray, K. (2017). Dying is unexpectedly positive. *Psychological Science, 28*(7), 988–999. https://doi.org/10.1177/0956797617701186

Gray, N. S., Weidacker, K., & Snowden, R. J. (2019). Psychopathy and impulsivity: The relationship of psychopathy to different aspects of UPPS-P impulsivity. *Psychiatry Research, 272*, 474–482. https://doi.org/10.1016/j.psychres.2018.12.155

Harmon-Jones, E., & Allen, J. J. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology, 74*(5), 1310–1316. https://doi.org/10.1037/0022-3514.74.5.1310

Hicks, B. M., & Patrick, C. J. (2006). Psychopathy and negative emotionality: Analyses of suppressor effects reveal distinct relations with emotional distress, fearfulness, and anger-hostility. *Journal of Abnormal Psychology, 115*(2), 276–287. https://doi.org/10.1037/0021-843X.115.2.276

Hong, M., & Kleck, G. (2018). The short-term deterrent effect of executions: An analysis of daily homicide counts. *Crime & Delinquency, 64*(7), 939–970. https://doi.org/10.1177/0011128717719514

Huges, J., & Beiner, D. (2021) reghelper: Helper Functions for Regression Analysis. https://cran.r-project.org/web/packages/reghelper/index.html

Johnson, R. (1979). Under sentence of death: The psychology of death row confinement. *Law & Psychology Review, 5*, 141–193.

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition & Emotion, 13*(5), 467–480. https://doi.org/10.1080/026999399379140

Ko, E., Kwak, J., & Nelson-Becker, H. (2015). What constitutes a good and bad death?: Perspectives of homeless older adults. *Death Studies, 39*(7), 422–432. https://doi.org/10.1080/07481187.2014.958629

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kübler-Ross, E. (1973). *On death and dying*. Routledge.

Lasko, E. N., & Chester, D. S. (2020). What makes a "successful" psychopath? Longitudinal trajectories of offenders' antisocial behavior and impulse control as a function of psychopathy. *Personality Disorders: Theory, Research, and Treatment, 12*(3), 207–215. https://doi.org/10.1037/per0000421

Lasko, E. N., Chester, D. S., Martelli, A. M., West, S. J., & DeWall, C. N. (2019). An investigation of the relationship between psychopathy and greater gray matter density in lateral prefrontal cortex. *Personality Neuroscience*. https://doi.org/10.1017/pen.2019.8

Lawler, K. A., Younger, J. W., Piferi, R. L., Billington, E., Jobe, R., Edmondson, K., & Jones, W. H. (2003). A change of heart:

Cardiovascular correlates of forgiveness in response to interpersonal conflict. *Journal of Behavioral Medicine, 26*, 373–393. https://doi.org/10.1023/A:1025771716686

LeDoux, J. (2012). Rethinking the emotional brain. *Neuron, 73*(4), 653–676. https://doi.org/10.1016/j.neuron.2012.02.004

Levine, L. J. (1996). The anatomy of disappointment: A naturalistic test of appraisal models of sadness, anger, and hope. *Cognition & Emotion, 10*(4), 337–360. https://doi.org/10.1080/026999396380178

Lutz, J., & Krahé, B. (2018). Inducing sadness reduces anger-driven aggressive behavior: A situational approach to aggression control. *Psychology of Violence, 8*(3), 358–366. https://doi.org/10.1037/vio0000167

Marsh, A. A. (2013). What can we learn about emotion by studying psychopathy? *Frontiers in Human Neuroscience, 7*, Article 181. https://doi.org/10.3389/fnhum.2013.00181

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence, 29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Naldi, M. (2019). A review of sentiment computation methods with R packages. https://doi.org/10.48550/arXiv.1901.08319

Navarrete, C. B., & Soares, C. (2020). dominanceAnalysis: Dominance analysis for general, generalized and mixed linear models. R package version 2.00, https://cran.r-project.org/web/packages/dominanceanalysis

Newman, J. P., Curtin, J. J., Bertsch, J. D., & Baskin-Sommers, A. R. (2010). Attention moderates the fearlessness of psychopathic offenders. *Biological Psychiatry, 67*(1), 66–70. https://doi.org/10.1016/j.biopsych.2009.07.035

Ohbuchi, K.-I., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology, 56*(2), 219–227. https://doi.org/10.1037/0022-3514.56.2.219

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*(3), 277–293. https://doi.org/10.1037/0022-3514.89.3.277

Peterson, R. D., & Bailey, W. C. (1991). Felony murder and capital punishment: An examination of the deterrence question. *Criminology, 29*(3), 367–395. https://doi.org/10.1111/j.1745-9125.1991.tb01071.x

Pfattheicher, S., Lazarević, L. B., Westgate, E. C., & Schindler, S. (2021). On the relation of boredom and sadistic aggression. *Journal of Personality and Social Psychology, 121*(3), 573–600. https://doi.org/10.1037/pspi0000335

Plaks, J. E., Robinson, J. S., & Forbes, R. (2021). Anger and sadness as moral signals. *Social Psychological and Personality Science, 13*(2), 362–371. https://doi.org/10.1177/19485506211025909

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of emotion* (pp. 3–33). Academic Press. https://doi.org/10.1016/B978-012-558701-3.50007-7

Pond, R. S., Jr., DeWall, C. N., Lambert, N. M., Deckman, T., Bonser, I. M., & Fincham, F. D. (2012). Repulsed by violence: Disgust sensitivity buffers trait, behavioral, and daily aggression. *Journal of Personality and Social Psychology, 102*(1), 175–188. https://doi.org/10.1037/a0024296

Provoost, S., Ruwaard, J., Van Breda, W., Riper, H., & Bosse, T. (2019). Validating automated sentiment analysis of online cognitive behavioral therapy patient texts: an exploratory study. *Frontiers in Psychology, 10*, Article 1065, 1–12. https://doi.org/10.3389/fpsyg.2019.01065

Riediger, M., Schmiedek, F., Wagner, G. G., & Lindenberger, U. (2009). Seeking pleasure and seeking pain: Differences in pro-hedonic and contra-hedonic motivation from adolescence to old age. *Psychological Science, 20*(12), 1529–1535. https://doi.org/10.1111/j.1467-9280.2009.02473.x

Rinker, T. W. (2021). sentimentR: Calculate Text Polarity Sentiment. Version 2.9.0. https://github.com/trinker/sentimentr

Schumann, K. (2012). Does love mean never having to say you're sorry? Associations between relationship satisfaction, perceived apology sincerity, and forgiveness. *Journal of Social and Personal Relationships, 29*(7), 997–1010. https://doi.org/10.1177/0265407512448277

Schumann, K. (2018). The psychology of offering an apology: Understanding the barriers to apologizing and how to overcome them. *Current Directions in Psychological Science, 27*(2), 74–78. https://doi.org/10.1177/0963721417741709

Schumann, K., & Dragotta, A. (2021). Empathy as a predictor of high-quality interpersonal apologies. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2786

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: Satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology, 94*(1), 116–132. https://doi.org/10.1037/0022-3514.94.1.116

Söderberg, P., & Björkqvist, K. (2020). Victimization from peer aggression and/or bullying: Prevalence, overlap, and psychosocial characteristics. *Journal of Aggression, Maltreatment & Trauma, 29*(2), 131–147. https://doi.org/10.1080/10926771.2019.1570410

Spice, A., Viljoen, J. L., Douglas, K. S., & Hart, S. D. (2015). Remorse, psychopathology, and psychopathy among adolescent offenders. *Law and Human Behavior, 39*(5), 451–462. https://doi.org/10.1037/lhb0000137

Staub, E., Pearlman, L. A., Gubin, A., & Hagengimana, A. (2005). Healing, reconciliation, forgiving and the prevention of violence after genocide or mass killing: An intervention and its experimental evaluation in Rwanda. *Journal of Social and Clinical Psychology, 24*(3), 297–334. https://doi.org/10.1521/jscp.24.3.297.65617

Sylvers, P. D., Brennan, P. A., & Lilienfeld, S. O. (2011). Psychopathic traits and preattentive threat processing in children: A novel test of the fearlessness hypothesis. *Psychological Science, 22*(10), 1280–1287. https://doi.org/10.1177/0956797611420730

Thomson, N. D., Kevorkian, S., Blair, J., Farrell, A., West, S. J., & Bjork, J. M. (2021). Psychophysiological underpinnings of proactive and reactive aggression in young men and women. *Physiology & Behavior, 242*(113601), 1–10. https://doi.org/10.1016/j.physbeh.2021.113601

Umbreit, M. S., Coates, R. B., & Vos, B. (2006). Victim offender mediation: An evolving evidence-based practice. Handbook of restorative justice: A global perspective, 52–62.

Vansteelandt, K., & Van Mechelen, I. (2006). Individual differences in anger and sadness: In pursuit of active situational features and psychological processes. *Journal of Personality, 74*(3), 871–910. https://doi.org/10.1111/j.1467-6494.2006.00395.x

Villar, G., Arciuli, J., & Paterson, H. M. (2014). Remorse in oral and handwritten false confessions. *Legal and Criminological Psychology, 19*(2), 255–269. https://doi.org/10.1111/lcrp.12012

West, S. J., & Chester, D. S. (2021). The Tangled Webs We Wreak: Examining the structure of aggressive personality using psychometric networks. *Journal of Personality*. https://doi.org/10.1111/jopy.12695

Zechmeister, J. S., & Romero, C. (2002). Victim and offender accounts of interpersonal conflict: Autobiographical narratives of forgiveness and unforgiveness. *Journal of Personality and Social Psychology, 82*(4), 675–686. https://doi.org/10.1037/0022-3514.82.4.675

Zhan, J., Ren, J., Fan, J., & Luo, J. (2015). Distinctive effects of fear and sadness induction on anger and aggressive behavior. *Frontiers in Psychology, 6*, Article 725, 1–12. https://doi.org/10.3389/fpsyg.2015.00725

Zimring, F. E. (2004). *The contradictions of American capital punishment*. Oxford University Press.