COMPUTERIZED TEXT-ANALYSIS OF OFFENDERS OF MASS SHOOTINGS: AN INVESTIGATION OF MORAL FOUNDATIONS USING LINGUISTIC INQUIRY AND WORD COUNT

A THESIS

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

MASTER OF SCIENCE

BY

KY BRAY

DR. THOMAS HOLTGRAVES – ADVISOR

BALL STATE UNIVERSITY

MUNCIE, INDIANA

JULY 2022

Abstract

The present study used a computational linguistic approach to examine moral foundations, emotionality, and personal concerns in the written communications of mass shooters. Writings by mass shooters (N = 36) were harvested from an online database and coded for writing type using five writing categories: manifestos, blog posts, social media, private letters, and journal entries. They were then submitted to linguistic analyses using Linguistic Inquiry and Word Count (LIWC). Shooters' writings were then compared to a sample of prisoners (N = 35) who were convicted of violent crime along with the normative data available in LIWC. The Moral Foundations Dictionary (MFD) was also used to predict differences between the samples, although no significant differences were found between shooters and prisoners on the five moral foundations. Overall, results of these analyses indicated that mass shooters primarily use high rates of negative emotion and swear words. Unlike previous studies, however, mass shooters were comparatively low on cognitive processes and complexity relative to the prisoner sample. Exploratory analyses were then conducted using the entire LIWC2015 and MFD dictionaries to identify the set of word categories that maximally predicts differences between groups across writing types.

Keywords: linguistic inquiry and word count, mass shootings, natural language use, moral foundations theory

Acknowledgements

Foremost, I want to thank Dr. Thomas Holtgraves for his guidance and mentorship throughout my time as a graduate student. Without him, the completion of this study would not have been possible. I also recognize Dr. Alexandria Johnson and Dr. W. Holmes Finch for their patience and expertise. Above all else, I want to thank all of these individuals for their time, which I too often did not appreciate.

Most of all, thank you to Kelsey, my first reader and best friend.

Table of Contents

Introduction

In many classical works of literature, authors explore perspectives of vengeful or nihilistic individuals who are driven to violence. Characters like Mephistopheles (Goethe, 1808/1985), Lucifer (Milton, 1968), and Stavrogin (Dostoyevsky, 1871/2008) have provided scholars content for analysis across centuries—perhaps millennia. Similarly, the real-world comparison of this pathology is often examined qualitatively by psychologists and criminologists. Some have argued that examining *legacy tokens*, written or spoken artifacts that foreshadow a pending mass murder event, can provide insight that may predict or classify offenders of mass shootings (Langman, 2009; Langman, 2015). Indeed, some individuals who commit mass murder write journals or manifestos about their intentions, and many are active in conveying their grievances online prior to an event. As such, the value of these communications is implicit in their relationship to the perpetrator and a mass murder event. Examining these communications is then likely to provide some insight to the common cognitive processes across mass shooters.

However, there is little consistency in the methods used to extract information from mass shooters' communications, and there is much disagreement regarding what exactly is learned from them. For example, some have merely attempted to describe and classify mass shooters. Langman (2009) and others (*e.g.,* Dutton et al., 2013) employed qualitative methods that use the content of mass shooters' communications and life-history to sort them into diagnostic categories based on the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; APA, 2013). Similarly, others have used software, such as Psychiatric Content Analysis and Diagnosis (PCAD), to test for rhetorical devices indicative of toxic masculinity and a variety of psychosocial issues related to psychiatric disorders (*e.g.*, anxiety, hostility, psychosis, quality of life, etc.; Sapru, 2019). Comparatively, Neuman et al. (2015) were more ambitious. They used an automated text-analysis software to validate predictive models of the personality traits of mass shooters.

These studies can be viewed as examining what McAdams (1995) termed the first level of personality: traits. In other words, the intended outcome from these studies is concerned with identifying characteristics of mass shooters that allow for trait-based description or shooter predictions. As such, the present study views these types of studies as either useless in scientific discourse or as stand-alone works that offer few options for future empirical research. However, people's communications may better reflect the second and third levels of personality as described by McAdams (Chung & Pennebaker, 2018). McAdams (1995) proposed that the second level of personality consisted of personal concerns such as motivation, development, and strategy. Compared to traits, McAdams argued that personal concerns describe what people want, what they do not want, and what they are willing to do to (1) avoid things they do not want and (2) attain things they do want. Personal concerns therefore allow for individuals' characteristics to be couched inside temporal-spatial contexts, and this is the level of personality with which the present proposal is concerned.

It is worth mentioning that McAdams (1995) described a third and deeper level of personality that is conveyed through the stories people maintain about themselves. McAdams called this life story *identity*, and it rests on the assumption that people produce self-narratives that integrate their experiences across the life-span. For example, when writing a life story, people use narrative arcs consisting of themes like agency, communion, or redemption (McAdams, 2006).

While interesting, exploring the identities of mass shooters is likely beyond the possibilities of this study, because most of these individuals did not record their life stories. Instead, the few who did leave behind written artifacts, communicated their emotional responses and future intentions in much shorter and context specific communications (Langman, 2009a). These are likely better considered a reflection of personal concerns. As such, the present proposal argues that computerized text-analyses can provide interesting and valuable insight into the psychosocial processes of mass shooters through their communication style. From this view, shooters' writings are considered one potential window through which premeditated violent behavior can be studied. In the present study, I collected the largest sample of writings by mass shooters used in a computerized linguistic analysis to date. They were then analyzed using Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2015) and compared to a sample of prisoners convicted for various forms of violent crime.

"Mass Shooter" Defined

Mass shootings are known by many names including school shootings (Langman, 2015), spree killings (Federal Bureau of Investigation [FBI], 2008), active shooting events (Kelly, 2010), rampage shooters (Lankford, 2013), pseudo commandos (Kop et al., 2019) and many others. Although there are some differences between the terms, most of these use the same operational definition based on four components: (1) a gun is used to shoot multiple people; (2) the shooting occurs within a 24-hour period; (3) the shooting is in a public location; (4) there is no evidence of other criminal activity (*e.g.*, robbery, drug related violence) that led to the shooting. While mass shootings are a subcategory of other forms of mass murder, the perpetrators of mass shootings have unique features that differentiate them from others. Duwe (2007; 2020) examined the patterns of different types of mass violence by comparing prevalence, casualty rates, and offender characteristics. They found that mass murder events (excluding mass shootings) were responsible for an average of 5.4 deaths and 4.0 casualties, compared to mass shootings which had an average of 7.21 deaths and 10.77 casualties. Accordingly, mass shootings are much more rare than other forms of mass murder in the United States. In the 20th century alone (1900-1999), there were 909 mass murders in the U.S. (9.18 per year), compared to 158 mass shootings between 1976 and 2018 (3.76 per year).

Mass shooters are also quite different from perpetrators of other forms of mass murder. Approximately 97% of all mass shootings are carried out by a lone offender, and 80% of mass shooters were under the age of 45 with an average age of 35 (Duwe, 2020). Furthermore, these individuals appear to spend a great deal of time planning their attack. In 2007, Duwe found that mass shooters (relative to other mass murder) are more likely to target strangers, and 57% of them died by suicide or suicide by police. They also demonstrated that the rate of suicidal behavior in mass shooters is more than 10 times higher than other forms of mass murder.

Most research focuses on the demographic or descriptive similarities between mass shooters that may foretell their behavior. These studies have identified various psychosocial risk factors such as difficult temperament, social marginalization, or history of aggression (Newman et al., 2004; O'Neill et al., 2016). McGee and DeBernardo (1999) examined the behavioral profiles of individuals responsible for 12 school shootings occurring in the U.S. during the 1990's. They found that most school shooters had very few friends and were rated as being immature with poor social skills. McGee and DeBernado also found that school shooters had cognitive abilities within the normal range, although they tend to have an unstable sense of self and are often confused about sexual issues. Prior to their attack on a school, the shooters tended to experience some form of psychosocial distress such as humiliation, rejection, or loss.

In a more recent study, Lankford (2013) conducted a comparative analysis of 81 terrorists who died by suicide and mass shooters. He found that the characteristic differences between mass shooters and terrorists who died by suicide were almost nonexistent. Instead, both had many of the same problems and stressors including social marginalization, family dissent, work or school related issues, and ongoing or precipitating crises. According to Lankford, these individuals often cite negative social interactions or foiled achievement of their goals as a source of rage (Lankford, 2015). Newman and Fox (2009) similarly found that social marginalization and individual differences in psychosocial risk factors (*e.g.*, desperation, suicidal ideation, familial problems) were among the most important predisposing features of individuals involved in mass shootings.

Newman and Fox (2009) argued that mass shooters' extreme negative emotionality—potentially caused by the psychosocial problems described above—often exacerbates preexisting mental health problems. Indeed, according to Duwe (2020), approximately 60% of all mass shooters have evidence of a pre-existing mental health problem. However, these studies only describe some of the commonalities between shooters. Given that approximately 50% of all mass shooters communicated their intentions prior to the attack (Lankford, 2013), other approaches have examined the content of these communications in relation to their psychosocial distress.

Characteristics Based on Written Artifacts

Peter Langman is perhaps one of the most well-known researchers to investigate the written artifacts of mass shooters, and he is responsible for compiling one of the most complete databases of documents related to mass shootings (schoolshootings.info) In 2009 (2009a; 2009b), he created a typology of mass shooters that integrated their writings and psychosocial risk factors into a classification system based on psychiatric diagnostic criteria. In developing his typology, Langman collected shooters' communications, court documents, police reports, and psychiatric assessments. He then examined their writing style and content within the context of reports from multiple sources, and used his examination to create a mass shooter typology based on the *DSM-IV* (American Psychiatric Association [APA], 1994). This classification system consisted of three types of individuals: traumatized, psychotic, psychopathic. According to his typology, traumatized shooters experienced physical and/or sexual abuse; psychotic shooters showed symptoms of schizophrenia or schizotypal personality disorder; and, psychopathic shooters demonstrated narcissism, lack of empathy, and sadistic behavior. Langman (2009a; 2009b) then applied this typology in idiographic assessments of 10 mass shooters treating the categories as mutually exclusive (2 psychopathic, 5 psychotic, 3 traumatized). This was applied to another 48 mass shooters in 2015, wherein Langman allowed the shooters to fall under multiple categories at once (Langman, 2015).

Despite the improvement in the classification system by allowing for heterogeneity among the shooter categories, Langman's (2015) typology has a few logical problems. Aside from the above differences in the categories, Langman (2009a) argued that the two defining features of offenders were rage and desire for revenge. Furthermore, as in previous descriptions (*e.g.,* Lankford, 2015), these features were likely caused by an overwhelming sense of victimization or humiliation (real or perceived) which eventually led to an attack on the perceived origin of their negative emotion (Langman, 2015). Considering these commonalities across the shooter categories, traumatized shooters perhaps should be the most prevalent. Yet, instead, traumatized shooters are the least prevalent, and the most common category among the idiographic assessments conducted by Langman (2009a; 2009b; 2015; 2016) is the psychotic category. Nonetheless, Langman (2016) argued that there was *substantial* evidence for paranoia, grandiose delusions, auditory hallucinations, and disorganized thoughts in the population.

The latter point made by Langman (2016) is suspect given that most evidence of psychosis symptoms were derived from shooters' journals. Dutton et al. (2013) also argued against Langman's typology due to construct validity issues in his conceptualization of psychopathy. That is, Langman (2009a; 2015; 2016) focuses heavily on callousness while largely ignoring other traits associated with psychopathy. For example, Langman disregards evidence suggesting those higher on psychopathic traits are not commonly concerned with anxiety, rejection, or few emotional bonds (Patrick, 2007). Because of this, Dutton et al. (2013) re-examined the written communications of four mass shooters using Langman's methodology. Similar to previous research, they found that these individuals were fixated on rejection by others and had a tendency to over-estimate the degree to which they received negative treatment. According to their interpretation, Dutton et al. (2013) found that the mass shooters whom they examined met criteria for paranoid personality disorder with narcissistic tendencies.

Others have conducted similar studies using quantitative techniques to find support for categorical diagnoses of mass shooters. Sapru (2019) used the Gottschalk-Gleser content analysis method, which is a technique that categorizes words and attitudes associated with psychological dimensions as they are used in natural speech (Gottschalk & Gleser, 1969). This was paired with the Psychiatric Content Analysis and Diagnosis (PCAD)—a software that tests for anxiety, hostility, thought distortions, depression, and a variety of other text-based indicators of psychological well-being (GB Software, 2016). In idiographic assessments of three mass shooters, Sapru found evidence for thought distortions, social alienation (designed to measure isolation tendencies in schizophrenia), and outward hostility. These results appear to support

Langman's (2016) findings regarding the prevalence of psychosis among the mass shooter population.

Neuman et al. (2015) found contradictory results. They compared the communications by six shooters (one text each) to a comparison group of 6,056 texts using automatic text analysis and vectorial semantics to identify co-occurring words. They then measured the similarity between the shooters' texts and the comparison group based on words usage relating to four personality types (developed from DSM-V criteria): depression, paranoid personality disorder, narcissistic personality disorder, schizotypal personality disorder. Unlike previous studies, their results provided support for narcissistic personality disorder, although they did find evidence of trait revengefulness which has been demonstrated elsewhere.

While Langman's (2009a) typology provides an interesting interpretation of mass shooters' psychology, it does not allow for psychological comparisons to be made empirically. Likewise, the problem with qualitative assessments (*i.e.*, Dutton et al., 2013; Langman, 2009a; Langman, 2015) of shooter's communications is that they are subject to the interpretations of the rater. They do not use comparison groups and are assessed by a rater(s) who know how the document is related to a person and event. This makes it difficult to determine whether the typologies above can distinguish between shooters' communications and other texts, or whether they identify any substantial contrasts between violent and non-violent individuals. Further, qualitative interpretations of these documents make it difficult to compare across groups or relate them to larger bodies of research that would contextualize their style or content.

The reliance on *DSM* (APA, 2013) categories to inform a *post hoc* interpretation of an event is also problematic across all of these studies. Foremost, this perpetuates a stereotype that mass shootings are common outcomes of mental illness, which does not accurately reflect reality. The shock value, casualty rates, and uniqueness of mass shootings has led the media, guns rights advocacy groups, and politicians to argue that mental illness is a causal factor (Metzl & MacLeish, 2015). Indeed, some reports indicate as many as 60% of mass shooters were diagnosed with or showed signs of mental health problems before their attack (Duwe, 2020). Yet, for another 40% of this population, there is no substantial evidence of psychopathology. Retrospectively diagnosing perpetrators of mass violence with a specific mental disorder is intellectually convenient and mostly based on conjecture. Baele (2014) identifies this as social coping behavior intended to alleviate society-wide shock following an event and argues that it is functionally irrelevant in scientific discourse. Statistically, the argument for mental illness as causal is also difficult to defend, given that the 12-month prevalence rate of mass shootings is three times higher than any mental disorder among adults (Duwe, 2020; Metzl & MacLeish, 2015).

Computerized Text Analyses of Personal Concerns

The above studies are examining mass shooter's communications from the wrong level of analysis. As described earlier, these communications are likely best suited for exploring what McAdams (1995) termed the second level of personality: personal concerns. This is due to the fact that shooters often describe their intentions, motivations, and frustrations through their communications. Of note, one study did examine the trait level of mass shooters with some success. Kop et al. (2019) used IBM Watson Personality Insights, an artificial intelligence software that used psycholinguistic predictive models to assess Big Five personality traits, to compare the personality profiles of 11 mass shooters to the general population. Their results demonstrated that the 11 mass shooters scored higher in Openness to Experience and lower in

Extraversion and Agreeableness relative to the general population, whereas there were no differences between samples in Neuroticism or Conscientiousness.

Aside from the results of Kop et al. (2019) the limitations and problems with studying the trait level suggest that the level of personal concerns is more compatible with studies of mass shooters communication. And, personal concerns can be examined using automated text-analysis software. Some examples of these have already been mentioned such as vectorial semantics. Others include Meaning Extraction Method (MEM, Chung & Pennebaker, 2008)—a factor analytic strategy that quantifies content domains in text—and IBM Watson Personality Insights (as used by Kop et al., 2019). Each of these examples are *open-dictionary* approaches to text-analysis, meaning that they do not rely on a pre-established lexicon to identify linguistic characteristics within a text. One area of concern with these methods is that they are more exploratory than closed content approaches and are sometimes highly individualized. In other words, there are few (if any) *a priori* hypotheses that can be made, because all indicators are derived atheoretically from texts. While open-vocabulary is valuable under certain contexts, the following review chooses to focus on *closed-dictionary* approaches.

Linguistic Inquiry and Word Count (LIWC), developed by Pennebaker and colleagues (2001), is one of the most widely used closed-dictionary, text-analysis programs. This software uses a pre-established dictionary to count percentages of words in a text based on more than 90 categories. The LIWC dictionary is hierarchical, and it can be sorted into two broader word categories that represent different psychological properties: content words and function words. Content words (nouns, verbs, adjectives, adverbs) show *what* a text is communicating by identifying *that which is acting*, *the action that is occurring*, or *that which is being acted upon* (Pennebaker et al., 2003; Tauczik & Pennebaker, 2010). Alternatively, function words (pronouns, articles, auxiliary verbs, adverbs, conjunctions, prepositions, and negations) show *how* the content is communicated and *that toward which attention is directed*. For example, the sentence, “I am happy about work,” contains three function words (I, am, about) and two content words (happy, work). The word, “I,” would be counted as a personal pronoun (more specifically a first-person singular pronoun), “am” would be counted as an auxiliary verb, and “about” would be counted as a preposition. Alternatively, “happy” would be counted as a positive emotion word, and “work” would be counted as a personal concern word (specifically under the work category).

LIWC, since its development, has proven to be useful in predicting a variety of psychological constructs and in demonstrating personal concerns related to those constructs. It has been used previously to extract Big Five personality dimensions from text. Those higher on extraversion tend to use positive emotion words (*e.g.,* love, nice), first person plural pronouns (*e.g.,* we, us), and second person pronouns (*e.g.,* you, yours), whereas those higher on neuroticism use more first-person singular pronouns (*e.g.,* I, me) and negative emotion words (*e.g.,* hurt, ugly; Ireland & Mehl, 2014). Theoretically, what this appears to indicate is that those higher on extraversion are concerned with related positive emotions to others and themselves (as indicated by “we”), whereas those higher are neuroticism are concerned with related negative emotions directly toward themselves.

This interpretation is based on the assumption that pronoun use is an index of attention. For example, individuals use significantly more first-person singular pronouns and fewer third-person pronouns (*e.g.,* they, them, theirs) when describing an event in which they were teased (except for men who always used more third-person plural; Kowalski, 2000). Conversely, they use more third person pronouns when describing a time in which they teased someone else. This premise can also be used to identify hierarchical status between people who are conversing. Sexton and Helmreich (2000) demonstrated this by analyzing the conversations between cockpit crew members in flight simulations at NASA. They found that rates of first-person plural pronouns increased linearly across the lifespan and with the rank of crew members. By contrast, crew members who were lower in rank used first-person singular pronouns at higher rates. Similarly, in experimental manipulations of social status, participants used more second person pronouns when they held higher status relative to the individual with whom they were conversing, and more first-person singular pronouns when they held lower status relative to their conversation partner (Kacewicz et al., 2009).

In studies of status or teasing, attention is directed toward those lower in status or to the victims of teasing—regardless of whom is speaking. This becomes especially interesting when paired with other word categories. For example, by analyzing suicide notes, individuals who died by suicide have been found to make more references to the self (first-person pronouns), death (*e.g.,* dead, grave), time (*e.g.,* hour), and religion (*e.g.,* God) than those who do not die by suicide (Handelman & Lester, 2007; Kim et al., 2019; Van den Nest et al., 2018). Similarly, those who died by suicide have shown to use less complex cognitive style than other writers as demonstrated by decreased prevalence in function word categories that organize thoughts (*i.e.,* articles & prepositions; Kim et al., 2019). Tauszcik and Pennebaker (2010) explain these findings by stating attention allocation and emotionality are implicit in the content and structure of utterances, and are therefore subject to personal and social situations. In the case of suicide notes, individuals appear to be focused on their own sense of suffering *only* (as demonstrated by first-person pronouns and low sentence complexity) in relation to the pending end to that suffering.

Stirman and Pennebaker (2001) tested this effect on 18 poets, because poetry has the highest rate of death by suicide among all art forms. They compared the poetry of poets who died by suicide to those who did not, and found that both groups used negative emotion, positive emotion, and death words at the same rate. And, the only difference between the language use of the two groups was use of self-references and collective references. That is, poets who died by suicide used more first-person singular pronouns and fewer first-person plural pronouns. Eichstaedt et al. (2018) similarly found that usage of first-person singular pronouns, negative emotion words, and depression symptom descriptors predicted future diagnoses of depression on Facebook.

High rates of emotion words, in the absence of pronoun use, can convey psychological immersion. Holmes et al. (2007) measured the degree to which language use may reflect cognitive and emotional processing of trauma in 25 domestic violence survivors, and whether this may predict symptom reduction over time. Over multiple writing sessions, reductions in physical pain (as measured by the Bodily Pain Scale) were associated with reductions in all affective words (positive and negative), although there was no relationship between writing style and symptoms of depression. Holmes et al. interpreted these results by considering emotional immersion in the traumatic event an index of the experience of physical pain.

Thus far, much of the literature reviewed regarding natural language use has focused on psychological distress. Indeed, this is one of the most widely researched areas of natural language use dating back to the initial publication of LIWC (Pennebaker et al., 2001). This research is informative to the present study considering the high rate of death by suicide and the psychological distress of mass shooters (Duwe, 2020). Yet, in light of the questionable proposed characteristics of mass shooters (*i.e.,* Langman, 2009a; Dutton et al., 2013), other works on narcissism, psychopathy, threat assessment are perhaps more directly related.

Personal Concerns in Narcissism, Psychopathy, & Threat Assessment

Carey et al. (2015) sought to test the commonly held belief that first-person singular pronoun use is a linguistic marker of narcissism. They administered five narcissism measures to 4,811 participants (15 different sample) and asked them respond to various open-ended prompts (*e.g.,* "Please describe yourself") for 15-minutes. Each sample received a different prompt that varied its medium (*e.g.*, expressive writing, interview, video-taped description). Across all prompts, narcissism was unrelated to the prevalence of first-person pronoun use. Holtzman et al. (2019) later re-examined the data from Carey et al. (2015) by expanding the word-categories of interest to examine other areas of potential interest to narcissism. Holtzman et al. (2019) found that narcissism was instead positively associated with rates of second-person pronouns and swear words. Narcissism also negatively predicted references to anxiety/fear (*e.g.,* worried, fearful), tentativeness (*e.g.,* maybe), and perception (*e.g.,* look, hear).

Hancock et al. (2013) conducted a similar study that compared the open-ended crime narratives of incarcerated violent offenders with psychopathy (measured by the Psychopathy Checklist-Revised [PCL-R]; N = 14) to a control sample of other prisoners (N = 38). Participant's responses were analyzed using Wmatrix (Rayson, 2008) which examines the affective tone of words and the Dictionary of Affect and Language (Whissell & Dewson, 1986)—a software that uses a pre-established lexicon to differentiate between context specific, semantic concepts. In their interviews, those who met PCL-R criteria for psychopathy used more cause/effect descriptors (*e.g.,* because, since) and more references to needs (*e.g.,* food, money). Participants in the psychopath sample also used more past tense and fewer present tense verbs, indicating more psychological distance from the crime itself. Further, psychopathy was predictive of fewer social references (*e.g.,* family), and greater difficulty in describing emotional events (as indicated by disfluencies such as “um”).

Computerized linguistic analyses have also been used to analyze communications of extremist organizations to predict the likelihood of a pending terrorist attack. By comparing communications from violent and non-violent extremist groups, Pennebaker (2011a) found violent extremist organizations used more personal pronouns and more affective words (positive and negative emotion) than the non-volent organizations. The violent group also showed marked decreases in complex thinking during the month prior to an attack. Similarly, leaders of terrorist organizations fluctuate in emotionality and complexity of thought depending on the situation (Pennebaker & Chung, 2007). That is, speeches and communications by leaders of terrorist organizations are more emotional and less cognitively complex in their expressions when a terrorist attack is about to occur. This was determined largely by increases in certainty words (*e.g.,* always, never) that support cognitive rigidity. Pennebaker and Chung (2007) note these results seem to indicate that embarking on a terrorist attack is an event that decreases complex thinking and increases the closeness of a group.

The results of terrorist threat assessment studies provide promise that LIWC or other text-analysis software could be used to investigate mass shooters’ communications. As discussed previously, a few studies have attempted this with the intention of predicting future mass shootings or classifying shooters’ mental health. Yet, only a few studies have utilized this approach with the intent of exploring shooter’s personal concerns. In one example, Kaati et al. (2016) used LIWC to analyze communications by 10 mass murderers. They found higher rates of negative emotion words and third-person pronouns among mass murder offenders compared to baseline archival data (provided with the LIWC software). The results from Kaati et al. also showed mass murderers are less concerned with friends (*i.e.*, buddy, neighbor) and preferentially focus on motivations like power (*i.e.*, superior, bully). Notably, their sample included individuals who were responsible for mass shootings such as Elliot Rodger, and those who were responsible for other forms of mass murder (*e.g.*, Ted Kaczynski).

Egnoto and Griffin (2016) also examined differences between suicide notes, LIWC expressive writing baseline, and 21 mass murder offenders. They found that mass murderers used more negative emotion words, particularly anger words, and fewer first-person pronouns than suicide notes. Mass murderers also made less references to the future (*e.g.*, will, soon), but they found there was no difference between the samples in positive emotion words or sexual references. Like Kaati et al. (2016), Egnoto and Griffin (2016) also did differentiate between offenders of mass shootings and offenders of other forms of mass violence.

Baele (2017) compared communications from 11 mass murderers—again, including both mass shootings and other types of mass murder (*i.e.*, bombing)—to speeches by three civil rights activists (Mahatma Ghandi, Martin Luther King Jr., Nelson Mandela) and the LIWC emotional writing baseline sample. Baele found that mass murderers used more negative emotion words (especially anger words), less positive emotion words, and similar rates of six-letter words and resentment compared to the civil-rights activists and LIWC baseline. Baele then compared murderers to civil rights activists in the proprietary category, *Analytic*, which is a LIWC category that indicates cognitive sophistication. Higher levels of Analytic scores have been demonstrated to predict collegiate level academic performance (Pennebaker et al., 2014). Interestingly, Baele (2017) found no significant difference between mass murderers and activists in Analytic scores, although both scored significantly higher on Analytic than the LIWC baseline.

Overall, the few computerized text-analysis studies conducted with mass shooters' communications, show high rates of negative affect, cognitive complexity, and third-person plural pronouns with low rates of positive affect, social words, and self-references. Based on previous LIWC research, this pattern of results indicates negative emotionality (anger), strong in-group versus out-group distinction (Pennebaker, 2011a), and complexly organized thoughts (Pennebaker et al., 2014). However, these results are somewhat difficult to interpret given the treatment of mass murderers as a homogenous sample and that samples often consist of mass murderers generally rather than differentiating between types of offenders. As mentioned previously, the rate of suicidal behavior in perpetrators of mass shootings is substantially higher than that of other types of mass murder, suggesting that mass shooters experience different psychosocial stressors (Duwe, 2020). Nonetheless, these few studies provide promise that there is value to be gained from examining the personal concerns of mass shooters, and their limitations indicate that there is room for expansion into other areas.

Moral Foundations Theory

One potential area of further inquiry into the personal concerns evidenced in the writings of mass shooters is offered by Moral Foundations Theory (MFT). MFT is an integrative theory that nests cultural explanations of human behavior within evolutionary psychology (Graham et al., 2013). It assumes that evolutionary processes required cultural responses to adaptive problems, and these responses have been imbedded within modern social structures and individual human perceptions. Haidt and Graham (2007) originally proposed MFT along with five moral foundations derived from universal human values and automatic emotional responses that appear cross-culturally. These are: Care/Harm, Fairness/Cheating, Ingroup/Loyalty, Authority/Subversion, Sanctity/Degradation. The Care/Harm foundation is an orientation to respond to visual and auditory signals of suffering. Fairness/Cheating involves concern for cheating or cooperation by individuals with whom one is interacting, and it ensures fairness of transaction. The Loyalty/Betrayal foundation organizes a person within their familiar and recognizable ingroup in relation to a despised, dangerous, or unknown “other.” Authority/Subversion weighs the importance of respect for the social hierarchy with the threat of tyranny. And, lastly, Sanctity/Degradation is an extension of the behavioral immune system which is concern with the purity of a person’s body, culture, or ingroup. Haidt (2012) later proposed a sixth moral foundation: Liberty/Oppression, which indicates concern in instances when one’s autonomy is being infringed upon.

Much of the research using MFT has been conducted comparing political ideologies. In perhaps the first study to do this, Graham et al. (2009) compared the moral foundations of 8,193 liberals and conservatives. Across multiple studies, liberals consistently endorsed Harm/Care and Fairness/Reciprocity at higher rates than the other three moral foundations. Comparatively, conservatives endorsed all five foundations almost equally. This pattern of results has been replicated multiple times, including by Graham et al. (2012) who obtained a nationally representative U.S. sample and Graham et al. (2011) who sampled from 11 different world religions. Graham et al. (2013) argue that this pattern of results is due to the way in which MFT addresses the third level of McAdams’ (1995) conceptualization of personality: *narrative identity*. That is, Graham et al. (2013) state that people’s life stories are often not fully articulated. Rather, these stories are often borrowed or impressed upon them by the ideologies and stereotypes of a culture.

MFT has proven to be useful, however, in contexts outside of political ideology such as moral character. Glenn et al. (2009) tested this in 2,517 adult volunteers by correlating responses to the Moral Foundations Questionnaire with the Levenson Psychopathy Scale. Those scoring higher on psychopathic traits indicated less concern about Care and Fairness with higher concern for Loyalty. However, higher scores on the psychopathy measure predicted willingness to violate all five of the foundations for money. Thus, individuals with higher levels of psychopathic traits appear to be more willing to violate morality in exchange for a tangible reward. These findings were replicated by Aharoni et al. (2011) who tested MFT in a sample of 222 male prisoners who were assessed for psychopathy. Again, psychopathy scores negatively predicted scores on Harm prevention and Hairness.

By contrast, Jonason et al. (2015) found that psychopathy (measured by the Dark Triad Dirty Dozen; Jonason & Webster, 2010) negatively predicted all moral foundations. In forensic sample of 219 detainees, Ye et al., (2021) also found that psychopathic traits negatively predicted all of the moral foundations. In 2016, Marshal et al. conducted a meta-analysis comparing different measures of morality in their relationships with psychopathy. Across six studies, psychopathic trats were negatively correlated with four of the five foundations, and there was no relationship between psychopathy scores and Loyalty.

At this juncture, it is worth noting that the mention of psychopathic traits within the context of MFT should not be conflated with the assumption that mass shooters may likewise show high levels of psychopathic traits. Yet, these results should be considered within the context of the inappropriateness of and lack of evidence for psychopathic characteristics among the mass shooter population (as identified above), despite typologies that may suggest otherwise (*e.g.,* Langman, 2009a).

Other psychological areas of interest within the MFT framework include perceptions of team sports rivalries in relation to team identification. Winegard and Deaner (2010) demonstrated that Loyalty scores predicted the extent to which participants identified with their favorite sports team. Alternatively, Tamborini et al. (2012) tested participants endorsement of MFT alongside their appreciation to violent media. They showed participants film clips that varied in levels of violent content and asked participants to assess whether they would enjoy the film. Participants who scored lower on Care rated violent scenes as being more interesting, whereas Fairness scores predicted enjoyment of film narrative containing justification for violence.

Of most interest to the present study, Graham et al. (2009) created a LIWC compatible, Moral Foundations Dictionary (MFD) to evaluate the prevalence of words in a text matching the five moral foundations. They then used this dictionary to compare moral foundations in sermons delivered at liberal or conservative churches. Their results showed that liberal religious leaders made more references to harm, fairness, and in-group, and conservative conversely made more references to authority and purity.

In a recent study, Kennedy et al. (2021) tested this dictionary in a large sample of social media users by correlating the MFD with scores on the Moral Foundations Questionnaire. They found reasonable support for the relationship between the dictionary and self-report measure, wherein Fairness concerns were the least well supported and Purity concerns were the most supported. This is likely due to the small size of the MFD, which contains only 295 words matching the MFT categories. As such, expansions to the MFD are currently in progress. For example, Araque et al. (2020) recently developed a larger lexicon from crowdsourcing assessments of MFT consisting of 1000 lemmas. Still, Chung and Pennebaker (2018) identified the MFD as one potential way to investigate personal concerns through language. Likewise, the

MFD is a likely and potentially promising area of expansion for further research into the communications of mass shooters.

Present Study

Past research on mass shooter communications is inhibited by many practical problems that limit the validity, reliability, and breadth of findings. Formal psychological assessments for mass shooters are often unavailable available or non-existent, and attaining a sufficient sample or comparison groups is difficult and time consuming. Further, existing studies of this type are limited insofar as they rely on qualitative interpretations that do not allow for comparisons to other samples and populations (*e.g.,* Langman, 2009a; 2009b; 2015). These studies also tend to assume all offenders of mass murder are a homogenous group (*e.g.,* Baele, 2017; Egnoto & Griffin, 2016) despite evidence suggesting there are important differences between them (Duwe, 2020). Therefore, the present study sought to address these issues by collecting a more select sample of mass shooters, using a new comparison group (incarcerated violent offenders), and introducing a new dictionary. The present study intends to accomplish this in a computational text-analysis of mass shooters and the prisoner sample using the Moral Foundations Dictionary (MFD; along with the standard LIWC2015 dictionary).

Since such computational analyses are few, much of this work remains exploratory. However, a few predictions can be made based on the literature reviewed.

Third Person Plural Pronouns. Mass shooters tend to have a heightened perception that they have been persecuted in some capacity (Langman, 2015). As such, all shooters were expected to use high rates of third person plural pronouns (*e.g.,* they, them) compared to the prisoner sample and the expressive writing baseline provided in LIWC2015 (Pennebaker et al., 2015). This is based on the findings from previous LIWC analyses (*e.g.*, Baele, 2017) and other research indicating men preferentially focus their attention on perpetrators of perceived or real transgressions (Kowalski, 2000).

Negative Emotion Words & Swear Words. Second, shooters' negative emotionality is one of their defining characteristics based according to qualitative studies (Langman, 2009a; Newman & Fox, 2009). More importantly, previous, quantitative studies (*e.g.*, Baele, 2017; Egnoto & Grffin, 2016) found that shooters used high rates of negative emotion words relative to comparison groups. Therefore, mass shooters were predicted to use more negative emotion words relative to the prisoner sample and the LIWC2015 baseline. Likewise, swear words are used to convey aggression and frustration (Tausczik & Pennebaker, 2010), and they are included as a negative emotion sub-category in LIWC2022 (Boyd et al., 2022). As such, shooters were also expected to use more swear words compared to the prisoner sample and the LIWC2015 baseline samples.

Analytic. Furthermore, past studies have suggested that mass shooters may tend to score high on the Analytic proprietary category (Baele, 2017). Analytic, which is computed by a combination of function word frequencies, measures the extent to which constructs and ideas are organized in a complex manner (Pennebaker et al., 2014). This is a relatively unique finding given that mass shooters also demonstrate heightened emotionality—a typically negative correlate of Analytic. Nonetheless, mass shooters were predicted to score higher on Analytic than the comparison groups.

Power. One of the unique features of mass shooters is the use of foiled achievement as a motivation for outrage (Lankford, 2015). Thus, it is likely that power is a salient personal concern for mass shooters. Kaati et al. (2016) also found that power words (*e.g.*, superior, bully) were used frequently compared to other word categories in idiographic assessments of 10 mass murderers. Mass shooters were then predicted to make more references to power than either the prisoner sample or the LIWC2015 expressive writing baseline.

Certainty. Pennebaker and Chung (2007) found that cognitive rigidity—as indexed by certainty words—was one of the characterizing features of terrorist communications. Kaati et al. (2016) also found that certainty words (*e.g.,* always, never) were preferentially used by mass murderers compared to other categories. Based on these studies, mass shooters were expected to use certainty words at higher rates than the comparison groups.

Moral Foundations Dictionary Predictions

Harm and Fairness. Lastly, a few predictions were made for the Moral Foundations Dictionary (Graham & Haidt, 2009). Relative to the comparison sample, shooters were predicted to score higher on references to Harm and Fairness. This was based on research suggesting that mass shooters are especially concerned with foiled achievements (Lankford, 2015), which may make harm and fairness salient moral issues for them. Furthermore, some mass shooters enjoy violent movies and video games (Langman, 2009; 2015), and higher Harm scores have been positively associated with enjoyment of violent media.

Loyalty. Mass shooters have been demonstrated to make few references to friends and affiliations (Kaati et al., 2016), and they often discuss social alienation as a source of their negative emotion (Newman & Fox, 2009). Therefore, mass shooters were expected to score lower on Loyalty relative to the prisoner sample.

Authority. It was slightly more difficult to make predictions for Authority and Purity. Authority, operationally indicates support for a social hierarchy, and descriptions of mass shooters suggest that many blame society for their adverse experiences (Langman, 2015;

Lankford, 2016; Newman & Fox, 2009). Therefore, shooters were expected to score lower on Authority, as they are likely to have a more negative view of the social hierarchy.

Purity. Lastly, some (McGee & DeBernardo, 1999) have suggested that mass shooters are preoccupied by sex. In Egnoto and Griffin (2016), mass murderers were not found to make more sexual references compared to the LIWC2015 expressive writing baseline or a sample of suicide notes. However, sexual references in some samples of suicide notes tend to be quite high relative to other groups (Stirman & Pennebaker, 2001). Likewise, Lankford (2013) suggested that many school shooters may be concerned with their lack of romantic success. Sexual (*e.g.*, horny, incest) words are likely to violate the Purity (*e.g.,* abstinence, virginity) foundation. Therefore, mass shooters were expected to make fewer references to Purity than the prisoner sample.

Method

Mass Shooter Sample

To test these hypotheses, shooters were harvested from Peter Langman's online database, (schoolshooters.info). The database contains legal records, police reports, mental health records, and mass shooter communications associated with 156 mass shootings between 1976 and 2018 (Duwe, 2020). Langman created this database by collecting court documents, law enforcement records, and other documents from local and national law enforcement agencies. These documents are often accompanied by a receipt of authenticity from the source (*i.e.,* FBI, public school district), and have proven to be a valuable resource in previous studies (*e.g.,* Baele, 2017; Egnoto & Griffin, 2016; Kaati et al., 2016; Neuman et al., 2015). Prior to data collection, the database contained at least one written artifact for 42 mass shooters, indicating that approximately 27% of shooters between 1976 and 2018 had usable data. Based on the sample sizes of previous studies on suicide notes (*e.g., N* = 18; Stirman & Pennebaker, 2001), domestic violence survivors (*e.g., N* = 25; Holmes et al., 2007), and mass shooters (*e.g., N* = 21; Egnoto & Griffin, 2016), the sample was expected to have sufficient power to detect differences between the samples.

All writings within this database were harvested based on the following criteria: (1) the author used a gun to kill one or more people in a public area in the absence of other criminal activity; (2) they have analyzable data such that their writings either exist in a digital format or can be reliably transcribed (*i.e.,* handwritten texts could be easily interpreted and converted to a digital format). Video content and transcriptions of spoken language use (*e.g.,* court transcripts) was excluded from analysis. Of note, there were six shooters for whom their writings could not be collected due to the difficulty of interpreting their handwritten journals. For the same reason, not all writings were able to be collected from each author who met the exclusion criterion. This procedure resulted in a final sample 36 shooters (60.6% White, all male) matching 194 separate writings. The shooters were between 12 and 41 years of age at the time of writing (*Mdn*. = 18.00, *M* = 20.08, *SD* = 5.07), and were of a similar age at the time in which they committed a mass shooting (*Mdn*. = 18.00, *M* = 20.60, *SD* = 4.77). A summary of racial demographics for this sample can be found in Table 1.

Prisoner Sample

Writings from the prisoner sample were harvested from the American Prison Writing Archive at Hamilton College (APW). These writings were collected and compiled to develop a database of the first-hand experiences of individuals who are incarcerated in the United States. To create the database, essays were collected from prisoners by soliciting requests for essays that describe what their life has been like since imprisonment or details regarding the events of their incarceration. The database holds approximately 2,100 essays with corresponding demographic information for each author.

To acquire a comparison group, essays by individuals within this dataset were selected using propensity matching for age, race, and type of crime (*i.e.*, violent) to the mass shooter sample. This was accomplished by using the search parameters on the APW website and identifying prisoners based on the aforementioned characteristics. That is, prisoners were first identified and matched by age at time of writing and race. This resulted in a relatively short list, as most authors within the database are older than 40 years of age. Prisoners were then selected if they committed a violent crime as defined by the Federal Bureau of Investigation (2010): murder, nonnegligent manslaughter, forcible rape, robbery, aggravated assault. Once identified, their writings were then compiled into the dataset alongside the shooter writings. This resulted in writings by 36 individuals (63.9% White, all male) ranging in age from 14-40 (*Mdn*. = 23, *M* = 23.93, *SD* = 5.99). A summary of racial demographics for the sample is shown in Table 1.

It is worth noting that there were a number of limitations to this approach. Foremost, the APW database contained few prisoners of Asian, Latinx, or Native American descent who were within the desired age range. Furthermore, the database contained predominately Black authors, whereas only 1.5% of the shooters were Black. Lastly, prisoners age at the time of incident (*i.e.*, the date of their crime), was concealed, and only age at time of writing was available. Thus, it was not possible to perfectly match the prisoner sample demographics to that of the shooters. Nonetheless, despite these issues, the selection procedure was able to reasonably match prisoners to shooters on relevant demographic characteristics.

LIWC2015 Baseline Comparison Samples

To provide context to the LIWC2015 analyses, mass shooters' communications were compared to the expressive writing and blog post word usage summary statistics provided with the *LIWC2015 Language Manual* (Pennebaker et al., 2015). The expressive writing summary statistics were derived from 29 samples with more than 2,500 authors in studies in which participants were randomly assigned to write about emotional topics (*e.g.*, traumatic experiences). Participants in this sample are from varying demographic and age populations (*e.g.*, elementary students, college students, prisoners, elderly), and base rates for all LIWC2015 categories are provided in the manual. Similarly, the blog post summary statistics were collected from multiple studies, and consist of ~37,000 blogs posted by different individuals.

Communication Coding Procedure

All harvested data were given an identification number corresponding to a specific author and placed within a dataset along with all of the other harvested writings. Variables within the dataset were created for demographic information related to each author including: age at time of attack, age at time of writing, race, and sample identifiers (*i.e.*, mass shooter = 1, prisoner = 0). All communications were also coded using two, broad, hierarchical communication types: *public*, denoting a communication made to a public audience; and *private*, denoting a communication made to oneself (*i.e.*, journal entry) or to a single individual (*i.e.*, letters). These broad categories were further coded into more specific categories, such that the *public* category consisted of blog posts, manifestos/public letters, and social media posts. The *private* category consisted of journal entries and letters/emails sent to specific individuals. Using this procedure, every communication within the dataset had two variables indicating communication type: one that indicates the broader communication family (public or private) and one that indicates one of the five, specific types of communication within that family.

The immediate intention of this classification system was to control for the medium and audience of shooters' communications. This is largely due to the ways in which communication styles are likely to vary depending on the audience with whom the individual is speaking (Kacewicz et al., 2009) or depending on the platform on which the communication is made (for examples see the *LIWC2015 Language Manual* by Pennebaker et al., 2015). It is also important to note that the prompt given to the prisoner sample makes all of their communications fall within the *private, journal entry* communication category due to the similarity between the prompt they were given and those used in expressive writing studies. This is a potential confound for two reasons. First, individuals within this sample were asked to describe their experiences in prison and the events leading up to their arrest to an external audience. As such, it is likely that these writings are most similar in style to online journal entries in the mass shooter's sample, whereby individuals publicly respond to various prompts online. Therefore, it could be argued that all of the prisoners' writings should be coded as public, blog posts. Second, the homogeneity of communication type among the prisoner sample is a potential confound in omnibus tests comparing across writing styles, because the shooter sample contains all types of writings. Lastly, not all shooters had a writing sample matching each of the categories and subcategories of communication type. This represents yet another confound in the data, as shooters with more communications are likely to be disproportionately represented across communication categories. However, it was deemed necessary to separate multiple writings by the same author, because they often were written for different purposes, audiences, and during different times. For these reasons, communication types were submitted to exploratory follow up analyses to examine type by sample interactions and within group differences for the mass shooter sample.

In the shooter sample, the writing- type coding scheme resulted in 102 (52.6%) public writings and 92 (47.4%) private writings. The largest, lower-level categories were public blog posts (*N* = 76, 39.2%) and journal entries (*N* = 84, 43.3%), followed by public manifestos (*N* = 20, 10.3%), private letters (*N* = 8, 4.1%), and social media posts (*N* = 6, 3.1%). Descriptive statistics across sample and writing type are presented for all LIWC2015 categories in Table 6.

Software

It is important to note that a dictionary or software other than LIWC2015 could have been used. A variety of open-vocabulary approaches are available such as Meaning Extraction Method (MEM; Chung & Pennebaker, 2008). Other studies (*e.g.,* Sapru, 2019) have used content analyses that are designed to detect forms of psychopathology as well. Other closed content dictionaries were also considered, such as the Dictionary of Affect in Language (DAL; Whissell & Dewson, 1986) that examines the affective tone of words, or Wmatrix (Rayson, 2008) which applies a pre-established lexicon to differentiate and detect context specific, semantic concepts (*i.e.,* "*fly* away" versus "the *fly*"). Nonetheless, despite the limitations of LIWC, it seemed suitable for the present study, as it allows for comparisons against normative data and has previously proven useful in studies on personal concerns (Chung & Pennebaker, 2018).

Linguistic Inquiry and Word Count

Linguistic inquiry and word count (LIWC2015; Pennebaker et al., 2015) was used to analyze all texts once they were harvested. LIWC2015 is a closed-content, text analysis software that counts word percentages based on a pre-established dictionary. The dictionary contains more than 90 categories that are psychological (*e.g.,* cognitive, social, emotional), content (*e.g.,* death, religion, achievement), or grammatical (*e.g.*, articles, prepositions, pronouns). The dictionary is hierarchical and each of the three broad categories are broken down into multiple, more specific word-level categories. For example, the grammatical category "pronouns" is further divided into first-person (*e.g.*, I, me, my), second-person (*e.g.*, they, them), third-person (*e.g.*, we, us), and impersonal (*e.g.*, all, it, many) pronouns. Other LIWC2015 computations include four proprietary language variables (Clout, Analytic, Authenticity, Tone) and three general descriptor categories (words per sentence [WPS], percent of target words captured [Dic], percent of words longer than six letters [Sixltr]).

The percentage of words captured within a text (Dic) can be viewed as a reliability estimate for closed-content linguistic analyses. This feature allows researchers to judge the extent to which a variance within a text has been adequately captured by the dictionary and determine whether a larger or alternative dictionary is needed. In the present study, approximately 82.80% (SD = 10.16) of words in the writing samples matched the dictionary, suggested that LIWC2015 captured most of the variance in word usage. A summary of descriptive statistics for the rates of words matching the LIWC2015 dictionary across the two samples can be found in Table 2.

LIWC Dictionary Development. In the development of the LIWC dictionary, the words belonging to each category were generated based on common grammatical rules and conceptual groupings of words (Pennebaker et al., 2001). Words and word categories were added to the LIWC dictionary by consulting standard English dictionaries, thesauruses, and psychological test instruments (*e.g.*, emotional rating scales; Pennebaker et al., 2015). The categories were then validated by at least three out of four judges who confirmed words as belonging to the category. When judges did not agree on the inclusion of a word to a specific category, online and other sources were consulted to determine a word’s common usage and meaning. Words for which there was no conclusive agreement were omitted from the dictionary.

Next, words were quantified as a percentage of total words from almost 200,000 text files. Words were treated as item responses and used to compute internal consistency of word categories. Categories with low base rates, low internal reliability, or infrequent use by researchers were removed. The dictionary has been updated multiple times since the initial version (*e.g.*, Pennebaker et al., 2007; Pennebaker et al., 2015) by removing categories with low base rates and adding, splitting, or combining categories as rated by judges (for a review see Tauzcik & Pennebaker, 2010).

Analytic Category. One comparison of interest is the complexity of thought with which mass shooters express themselves in their writings. In Baele (2017), mass murderers were demonstrated to score especially high on this category relative to the expressive writing baseline and a comparison group of civil rights activists. Analytic is a proprietary LIWC2015 category that uses eight, word categories: personal pronouns, impersonal pronouns, auxiliary verbs, articles, prepositions, conjunctions, negations, and adverbs. In total, the eight categories account for 370 function words (Pennebaker et al., 2014). The formula is as follows: Analytic = 30 + articles + prepositions – personal pronouns – impersonal pronouns – auxiliary verbs – conjunctions – adverb – negations.

Moral Foundations Dictionary. Graham and Haidt (2009) created the LIWC2015 compatible, Moral Foundations Dictionary (MFD) by developing associations, synonyms, and antonyms for the base words of each foundation, including full words and word stems. This resulted in a list of words that affirmed and violated each foundation. Words were then sorted based on the closeness of their relationship with the foundations, and words that were either too similar or too different were removed from the dictionary. The remaining words were then tested using *LIWC2015*, and re-examined qualitatively to ensure that they accurately reflected moral foundations in context. The final dictionary consists of 295 words that either support or violate the five moral foundations (Harm, Fairness, Loyalty, Authority, Purity). That is, words that violate the foundation are negatively weighted, whereas those that support the foundation are positively weighted.

One notable limitation of the MFD is its size (Graham & Haidt, 2009). That is, the dictionary contains few words relative to the English lexicon, and base rates for the MFD categories tend to be quite low. Like the larger LIWC2015 dictionary, MFD allows researchers to extract the number of words within a text matching the dictionary. In the present study, approximately 2.17% (SD = 2.44) of the words used in either sample matched the dictionary. A summary of descriptive statistics for the rates of words matching the MFD dictionary across the two samples can be found in Table 2.

Procedure

Communications by the samples were collected, coded for author identification and relevant demographic characteristics. The communications were then assigned to writing type categories using the coding procedure outlined above, with the prisoner sample having only private, journal writings. Each available writing sample by shooters was also coded using the aforementioned criteria. For the shooter sample, a single communication by that author represented one case, and was coded for communication type based on the specific characteristics of that individual communication. Each communication and corresponding information were compiled into a single dataset which was submitted to LIWC2015 for analysis based on the 90+ LIWC categories and the 10 MFD categories.

Statistical Procedure

For all analyses, IBM SPSS Statistics 27.0.0.0 was used. First, a series of one-sample *t*-tests was used to compare mass shooters' communications to those of the *LIWC2015* expressive writing and blog post summary statistics. Since no summary statistics are available in the expressive writing sample for the Moral Foundations Dictionary, mass shooters' writings were compared individually to those of the normative data across only the original LIWC categories of interest.

Benjamini-Hochberg (1995) corrections were used to control for alpha inflation. As opposed to the Bonferroni method which controls for family wise error rate, the Benjamini-Hochberg method is less strict by controlling the false discovery rate (Haynes, 2013). This occurs by assigning ranks to *p*-values for all analyses, and computing a B-H critical value based on the number of tests, the rank of each *p*-value, and the false discovery rate (set at .05 in the present study). According to the Benjamini-Hochberg method, all *p*-values lower than their B-H critical value are considered significant. Following this procedure, only *p*-values less than .018 were considered statistically significant.

Two Poisson regressions were then used to determine whether the 10 *LIWC2015* word categories of interest are statistically useful in differentiating between the two groups. Poisson, as opposed to linear regression, automatically assumes that the words from each category are normally distributed. This procedure has been used in similar, previous studies (Kaati et al., 2016), and it has been noted that assuming a Poisson is justified for describing the number of occurrences of a certain word within a document (Ogura et al., 2013). In the first regression model, the samples were dummy coded and entered as outcome variables into the model, while the six *LIWC* categories were used as predictors. In the second, the samples were again used as outcome variables with the five MFT word categories entered as predictors.

Next, in an attempt to control for the potential confounds of writing type, analyses using the same procedure and all 10-word categories were conducted for private journal entries and for public blog posts. That is, public and non-journal (*i.e.*, letters) were excluded from these analyses. The word categories were then entered as predictors to examine whether they are statistically useful in differentiating between the groups and to identify the importance of individual predictors.

Lastly, exploratory analyses using forward, stepwise binary logistic regression models were conducted to examine the constellation of word categories that maximally differentiates between the prisoner and mass shooter samples. This consisted of entering all LIWC2015 and MFD word categories into the model as predictors and atheoretically extracting the combination of word frequencies that best predicts each group. In all analyses, entry was set to .05 and removal was set to .10. The five-level communication-type variable (consisting of social media, blog posts, manifestos, journal entries, private letters) was also used to in exploratory analyses to examine the effects of writing type. That is, exploratory, stepwise binary logistic regression models were conducted comparing shooters' blog posts and shooters' private journals to the writings of prisoners.

Results

Expressive Writing Baseline Comparisons

A series of one sample *t*-tests with were conducted to examine differences between shooter writings and the expressive writing summary statistics provided in the LIWC2015 manual (Pennebaker et al., 2015). After corrections for alpha inflation, these comparisons revealed six, significant differences between the shooter sample and the expressive writing baseline. As hypothesized, shooters (*M* = 49.82, *SD* = 26.97) scored significantly higher on Analytic than the baseline (*M* = 44.88, *SD* = n.a.), *t*(193) = 2.55, *p* = .011, *d* = .18, 95% CI [1.12, 8.77]. Shooters (*M* = .98, *SD* = 1.77) also used third person plural pronouns (*i.e.*, they) at a higher rate than the expressive writing baseline (*M* = .57, *SD* = n.a.), *t*(193) = 3.25, *p* = .001, *d* = .23, 95% CI [.16, .66], and shooters (*M* =2.00 , *SD* = 1.87) used more certainty words relative to the expressive writing baseline (*M* = 1.51, *SD* = n.a.), *t*(193) = 3.65, *p* < .001, *d* = .26, 95% CI [.22, .76]. As hypothesized, shooters (*M* = 2.56, *SD* = 2.58) made more references to power than the baseline (*M* = 2.02, *SD* = n.a.), *t*(193) = 2.94, *p* = .004, *d* = .21, 95% CI [.18, .91). Lastly, the largest differences between shooters and the expressive writing baseline was in negative emotion and swear word frequencies. That is, shooters (*M* = 3.99, *SD* = 3.25) made significantly more references to negative emotion relative to the expressive writing group (*M* = 2.12, *SD* = n.a.), *t*(193) = 8.01, *p* < .001, *d* = .58, 95% CI [1.41, 2.33]. Shooters (*M* = 1.32, *SD* = 2.26) also used more swear words compared to the expressive writing baseline *M* = .09, *SD* = n.a.), *t*(193) = 7.57, *p* < .001, *d* = .54, 95% CI [.97, 1.55].

The same comparisons were then conducted with one-sample *t*-tests while controlling for writing type, such that private, journal entries by shooters were compared to the expressive writing baseline. Negative emotion was again the largest difference between the two groups, with shooters (*M* = 4.20, *SD* = 2.71) using significantly more negative emotion words than the expressive writing group (*M* = 2.12, *SD* = n.a.), *t*(83) = 7.04, *p* < .001, *d* = .77, 95% CI [1.49, 2.67]. Shooters (*M* = 1.50, *SD* = 1.77) again used more swear word compared to the baseline (*M* = .09, *SD* = n.a.), *t*(83) = 7.24, *p* < .001, *d* = .79, 95% CI [1.01, 1.78]. Similarly, shooters (*M* = 2.22, *SD* = 1.71) used more certainty words than the expressive writing group (*M* = 1.51, *SD* =

n.a.), $t(83) = 3.80$, $p < .001$, $d = .41$, 95% CI [.34, 1.08], and shooters ($M = .91$, $SD = 1.14$) used third person plural pronouns at higher rates ($M = .57$, SD = n.a.), $t(83) = 2.72$, $p = .008$, $d = .30$, 95% CI [.09, .58]. However, when controlling for writing type, shooters ($M = 40.72$, $SD = 22.28$) did not score higher on Analytic ($M = 44.88$, SD = n.a.), $t(83) = -1.71$, $p = .091$, $d = -.19$, 95% CI [-8.99, .68]. Likewise, shooters ($M = 2.30$, $SD = 2.04$) also did not make more references to power ($M = 2.02$, SD = n.a.), $t(83) = 1.24$, $p = .218$, $d = .14$, 95% CI [-1.67, .72].

Blog Post Baseline Comparison

As with the expressive writing baseline, the entire sample of shooters was then compared to the blog post summary statistics using a series of one sample *t*-tests. Shooters ($M = 3.99$, $SD = 3.25$) again made more references to negative emotion than the blog posts ($M = 2.06$, SD = n.a.), $t(193) = 8.27$, $p < .001$, $d = .60$, 95% CI [1.47, 2.39], and they ($M = 1.32$, $SD = 2.26$) used more swear words ($M = .35$, SD = n.a.), $t(193) = 5.97$, $p < .001$, $d = .43$, 95% CI [1.47, 2.39]. Shooters ($M = 2.00$, $SD = 1.87$) used certainty words at a significantly higher rate compared to blog posts ($M = 1.56$, SD = n.a.), $t(193) = 3.28$, $p = .001$, $d = .24$, 95% CI [.18, .71], and shooters ($M = 2.56$, $SD = 2.57$) made more references to power than the blog posts ($M = 2.07$, SD = n.a.), $t(193) = 2.67$, $p = .008$, $d = .19$, 95% CI [.13, .86]. Lastly, shooters ($M = .98$, $SD = 1.77$) also used more third person plural pronouns compared to the blog posts ($M = .68$, SD = n.a.), $t(193) = 2.38$, $p = .018$, $d = .17$, 95% CI [.05, .55]. Interestingly, there was not a significant difference between shooters ($M = 49.82$, $SD = 26.98$) and blog posts in Analytic scores ($M = 49.89$, SD = n.a.), $t(193) = -.03$, $p = .973$, $d = -.002$, 95% CI [-3.89, 3.76].

The same comparisons were then conducted while controlling for writing type by including only shooter texts coded as public, blog posts. As expected, shooters ($M = 3.63$, $SD = 3.54$) again used negative emotion words at higher rates ($M = 2.06$, SD = n.a.), $t(75) = 3.89$, $p <$

.001, d = .45, 95% CI [.77, 2.38], and shooters (M = 62.63, SD = 26.10) were more Analytic (M = 49.89, SD = n.a.), $t(75) = 4.26$, $p < .001$, d = .49, 95% CI [6.78, 18.70]. Shooters' blogs (M = 1.49, SD = 2.95) again contained more swear words (M = .35, SD = n.a.), $t(75) = 3.38$, $p < .001$, d = .39, 95% CI [.47, 1.82.]. Shooters (M = 2.50, SD = 2.98) did not make more references to power (M = 2.50, SD = n.a.), $t(75) = 1.27$, $p = .207$, d = .15, 95% CI [-.25, 1.11], and shooters blogs (M = 1.08, SD = 2.59) did not use more third person plural pronouns than the LIWC2015 blogs (M = .68, SD = n.a.), $t(75) = 1.39$, $p = .168$, d = .16, 95% CI [-.17, .97]. Shooters blogs (M = 1.89, SD = 2.30) also did not use certainty words differently than the blog baseline (M = 1.56, SD = n.a.), $t(75) = 1.23$, $p = .221$, d = .14, 95% CI [-.20, .85].

Comparisons with Prisoner Sample

Two Poisson regression model were conducted to compare shooters to prisoners across the hypothesized word categories. First, the hypothesized LIWC2015 variables were entered to test whether they predicted group membership. Results of the omnibus test for this model indicated that the independent variables did not significantly predict differences in the groups, $X^2(228) = 66.05$, $p = 1.00$. Furthermore, model fit statistics suggested severe under-dispersion of the predictors (Pearson chi-square = .000) and thereby poor model fit. This was likely due to the low base rate of word counts and the high frequency of identically zero scores (as shown in Table 2). As such, the model was unable to estimate differences between the groups depending on rates of the word categories, and the results for this analysis were not further interpreted.

A similar problem was found when entering the five MFD categories as predictors. Model fit statistics were slightly better (Pearson chi-square = .02), although they still indicated under-dispersion. Similarly, the omnibus test for the model was not significant, $X^2(178) = 64.07$, $p = 1.00$. Like the first model, the results of this model were not able to be further interpreted.

Follow-up Comparisons with Prisoner Sample

Due to issues related to base rates of the hypothesized word categories, an alternative approach was used to ensure the absence of group differences. Two, binomial logistic regression models were conducted: (1) comparing differences between shooters and prisoners on the hypothesized LIWC2015 categories; (2) comparing differences between shooters and prisoners on the hypothesized MFD categories. In the first model, the Hosmer and Lemeshow test of model fit was not significant, $X^2(8) = 4.52$, $p = .807$, indicating that the model was a good fit for the data. The omnibus test also indicated that group membership was significantly predicted by the LIWC2015 word-categories of interest, $X^2(6) = 47.47$, $p < .001$, and the model explained 32.1% (Nagelkerke R^2) of the variance in group differences. Group membership was predicted with 86.1% accuracy by the model. As expected, swear words (*OR* = 2.75, 95%CI [1.10, 6.96]) were a significant predictor, with shooters using more swear words compared to prisoners. Two of the predictors were statistically significant, although opposite the hypothesized direction of the effects. Namely, prisoners scored significantly higher on Analytic (*OR* = .96, 95% CI [.94, .98]) and references to power (*OR* = .83, 95% CI [.72, .97]). Third person plural, negative emotion words, and certainty words were not significant predictors of group membership ($p > .05$). These results are presented in Table 3.

Next, a similar, binomial logistic regression model was conducted while entering the five MFD categories as predictors. According to the Hosmer and Lemeshow test, the model was not a good fit for the data, $X^2(7) = 33.31$, $p < .001$, and the omnibus test suggested that the predictors did not predict differences between the groups, $X^2(5) = 6.48$, $p = .262$, Nagelkerke $R^2 = .048$. Likewise, none of the MFD categories were individually significant predictors ($p > .05$). These results are also presented in Table 3.

Exploratory Analyses

Because prior research has not explored how the relationships between mass shooters and comparisons samples may vary as a function of writing type, analyses were conducted on the hypothesized LIWC2015 and MFD categories while using the writing type coding scheme as selection variables. Two, binary logistic regression models were conducted comparing shooters' journal entries (N = 84) to prisoners' writings. In the first model, the six LIWC2015 categories of interest were entered as predictor variables. The Hosmer and Lemeshow test indicated that the model was a good fit for the data, $X^2(8) = 7.08$, $p = .528$, and the omnibus test suggested that the six LIWC categories significantly predicted differences between the groups, $X^2(6) = 67.56$, $p < .001$, Nagelkerke $R^2 = .610$. The model also accurately predicted group membership with 70.0% accuracy. Similar to previous analyses, Analytic (OR = .94, 95% CI [.91, .97]) was significantly higher in the prisoner sample. Unexpectedly, prisoners also used more third person plural references (OR = .58, 95% CI [.34, .98]). None of the other categories (third person plural, negative emotion, certainty, swear) were significant predictors.

Shooters' journal entries were then compared to prisoners for the MFD categories in a binomial logistic regression. The Hosmer and Lemeshow test was significant, $X^2(7) = 20.07$, $p = .005$, and the model predicted group membership with 71.4% accuracy. The combination of MFD word categories did significantly predict group differences, $X^2(5) = 19.31$, $p = .002$, Nagelkerke $R^2 = .221$, although Authority (higher in prisoners) was the only significant predictor (OR = .11, 95% CI [.03, .43]).

Next, shooters' blog posts were compared to prisoners' writings across the six hypothesized LIWC2015 categories. Model fit was acceptable when entering the six, LIWC2015 categories as predictors, $X^2(8) = 13.67$, $p = .091$, and the overall model was significant, $X^2(5) =$

25.68, $p < .001$, Nagelkerke $R^2 = .286$. Furthermore, the model predicted group membership with 67.9% accuracy, although only two of the word-categories were significant predictors. Shooters scored significantly lower on references to power (*OR* = .82, 95% CI [.70, .98]), and used a higher rate of swear words (*OR* = 2.88, 95% CI [1.06, 7.83]). Analytic, third person plural pronouns, negative emotion, and certainty were not significant predictors.

Lastly, shooters blog posts were compared to prisoners using the MFD categories. The resulting model was a reasonable fit for the data, $X^2(8) = 12.04$, $p = .150$, although the overall model was not statistically significant, $X^2(5) = 4.37$, $p = .498$, Nagelkerke $R^2 = .051$. None of the MFD categories were significant predictors of group differences ($p > .05$). All of the results in this section can be found in Table 4.

Data Mining Exploratory Analyses

To explore the combination of word categories that maximally predicts differences between the prisoner and shooter writings, all LIWC2015 and MFD word categories (except frequency of dictionary words) were entered into a forward, stepwise binary logistic regression model using group as the outcome variable. The model ended in 13 steps, and the Hosmer and Lemeshow test indicated that the model was a reasonable fit for the data, $X^2(8) = 6.91$, $p = .547$. The resulting model was significant, $X^2(13) = 149.26$, $p < .001$, explained 82.3% (Nagelkerke R^2) of the variance in group differences, and correctly classified 96.5% of all cases (86.1% of prisoners and 98.5% of shooters). Thirteen, word categories were identified as significant predictors of group membership: word count (WC), first person singular pronouns (*e.g.*, I, me, my), third person singular pronouns (*e.g.*, he, she), prepositions, conjunctions, affect (*i.e.*, positive and negative emotion words), sad words (*e.g.*, crying, sad), friend words (*e.g.*, buddy, neighbor), sexual words (*e.g.*, horny, love), drives (*i.e.*, combination of affiliation, achievement, power, reward, risk), exclamation marks, and MoralityGeneral (combination of all MFD categories). Risk words (*e.g.*, danger, doubt) were also included in the model, although they were not a significant predictor (p = .055). Relative to prisoners, shooters used significantly more first-person singular pronouns (*OR* = 2.13, 95% CI [1.41, 3.22]), third person plural pronouns (*OR* = 4.19, 95% CI [1.96, 8.93]), emotion words (*OR* = 2.92, 95% CI [1.54, 5.51]), sexual words (*OR* = 53.104, 95% CI [3.43, 821.19]), friend words (*OR* = 25.12, 95% CI [3.07, 205.47]), exclamation marks (*OR* = 32.57, 95% CI [2.17, 487.77]), and overall references to morality (*i.e.*, MoralityGeneral; *OR* = 169.95, 95% CI [2.66, 10862.14]). Conversely, shooters used fewer overall words (*i.e.*, WC; *OR* = 1.00, 95% CI [.99, 1.00]), prepositions (OR = .40, 95% CI [.25, .65]) conjunctions (*OR* = .37, 95% CI [.20, .70]), sadness words (*OR* = .05, 95% CI [.01, .46]), drives (*OR* = .47, 95% CI [.30, .76]), and risk words (*OR* = .33, 95% CI [.11, 1.02]).

Next, to further explore how these relationships vary by writing type, two more stepwise binomial logistic regression models were conducted. First, only shooters writings coded as private journal entries (N = 84) were included. This model ended in nine steps and extracted seven predictors (WC, Clout, social words, sexual words, power word, home words, exclamation marks). The Hosmer and Lemeshow test was not significant, indicating that the model was likely a good fit for the data, $X^2(3) = .000$, $p = 1.000$. The omnibus test was statistically significant, $X^2(7) = 146.61$, $p < .001$, Nagelkerke $R^2 = 1.00$, and was able to predict group membership with 100% accuracy (100% of shooters and 100% of prisoners). However, the resulting model had a number of issues with local model fit. These included non-significant predictors and *very* small odds ratios for all predictors (*i.e., OR* < .001). Thus, to be conservative with the stepwise procedure and to avoid overfitting the data, the decision was made to use a simpler model identified in Step 6 (similar problems were found in Steps 7 and 8). Comparatively, Step 6, consisting of WC, Analytic, Clout, power, home, and exclamation marks, was also demonstrated to be a good fit for the data, $X^2(8) = 3.58$, p = .893. The omnibus test was significant, $X^2(6) = 116.81$, $p < .001$, Nagelkerke $R^2 = .882$, and the model predicted group membership with 94.2% accuracy (95.2% of shooters and 91.7% of prisoners). When writing a journal entry, shooters scored lower in total words (*OR* = .998, 95% CI [.997, 1.000]), Analytic (*OR* = .91, 95% CI [.85, .98]), Clout (*OR* = .95, 95% CI [.91, .99]), power (*OR* = .28, 95% CI [.11, .72]), and references to home (*OR* = .02, 95% CI [.002, .28]). Shooters, by comparison, scored higher than prisoners only in their use of exclamation marks (*OR* = 111.16, 95% CI [2.83, 4360.74]).

Lastly, considering that the prisoners' writings could have been coded as blog posts, a final, forward stepwise binomial logistic regression model was conducted comparing prisoners to only shooters' blog posts ($N = 76$). This model ended after 12 steps and suggested 10 predictors (prepositions, friend, causal words, drives, leisure, swear words, netspeak, colons, semicolons, and exclamation marks). This model was also statistically significant, $X^2(10) = 115.84$, $p < .001$, Nagelkerke $R^2 = .901$. However, similar to the previous analysis, the resulting model was determined to be a poor fit for the data based on odds ratios, predictor significance, and the Hosmer and Lemshow test, $X^2(7) = 25.72$, $p = .001$. Furthermore, this model did not substantially improve upon predictive power relative to previous, similar iterations. Step 11 had similar issues with model fit and predictive power, and the decision was mad to interpret Step 10. This model was determined to be a better fit for the data, $X^2(8) = 9.40$, $p = .310$, and accurately predicted 95.5% of cases (97.4% of shooters, 91.7% of prisoners). The omnibus test was significant, $X^2(8) = 108.55$, $p < .001$, Nagelkerke $R^2 = .868$, and consisted of eight categories: prepositions, friend, cause, drives, leisure, swear words, semicolons, and exclamation marks. In writing blog posts, shooters made more references to friends (*OR* = 16.33, 95% CI [.56, 478.50]), leisure (*OR* =

15.03, 95% CI [2.17, 104.10]), swear words (*OR* = 6.27, 95% CI [1.42, 27.68]), and used more exclamation marks (*OR* = 115.81, 95% CI [1.49, 8991.99]). Conversely, shooters used fewer prepositions (*OR* = .39, 95% CI [.21, .72]), causal words (*OR* = .18, 95% CI [.04, .74]), references to drives (*OR* = .34, 95% CI [.18, .64]), and semicolons (*OR* = .002, 95% CI [.000007, .55]). The results for each of the data mining analyses reported in this section are presented in Table 5.

Discussion

The goal of the present study was to further investigate the personal concerns of mass shooters by improving upon the sampling procedure used in previous studies. This was first accomplished by using a more selective inclusion criteria whereby all authors included in the sample used a gun to shoot one or more people in a public area in the absence of all other criminal activity. Furthermore, by coding the written artifacts into separate "writing type" categories, I was able to make more appropriate comparisons across separate writing contexts. Many of these comparisons were consistent with my hypotheses, especially those comparing shooters to the LIWC2015 expressive writing baseline. Alternatively, comparisons between mass shooters and the prisoner sample yielded some unexpected results, and exploratory analyses revealed a new and interesting constellation of salient personal concerns for mass shooters. As such, this study extends previous research on mass shooters and offers potential avenues for further investigation.

Emotionality

One of the most widely discussed characteristics of mass shooters is their emotionality (Dutton et al., 2013; Sapru, 2019; Langman 2009a; Langman 2009b; Langman, 2015; Newman & Fox, 2009). High negative affect and low positive affect was also demonstrated among mass shooters in linguistic analyses conducted by Baele (2017) and Egnoto and Griffin (2016). The present study appears to replicate these findings. Overall, mass shooters were found to use higher rates of negative emotion words relative to the LIWC2015 expressive writing baseline, the LIWC2015 blog baseline, and the prisoner sample.

Rates of other word categories suggest mass shooters were more preoccupied by their emotion relative to the other groups. In exploratory analyses, negative emotion words were not a predictor selected by any of the models. However, mass shooters did use significantly more many positive and negative emotion words (*i.e.*, affect) than prisoners when all writing types were included. In the same model, mass shooters also used two times as many first-person singular pronouns ($OR = 2.13$). This combination of pronoun use and emotionality is consistent with psychological immersion demonstrated in studies on suicide notes (Handelman & Lester, 2007), depression (Eichstaedt et al., 2018), and processing of trauma (Holmes et al., 2007). As hypothesized, shooters also used more swear words compared to both LIWC2015 baselines and prisoners. Swear words are often used informally to convey aggression and frustration (Tauzcik & Pennebaker, 2010).

One of the more novel findings from the exploratory analyses was the higher use of exclamation marks in the shooter sample. A conservative interpretation of this is that exclamation marks are a nuisance variable that were coincidentally selected in the stepwise process. Generally, punctuation is not commonly investigated nor discussed in computerized linguistic analyses, and many text-analysis software do not even include punctuation (Teh et al., 2015). However, the finding that exclamation marks consistently predicted mass shooters may be important to understanding their emotional state across writing types. Teh et al. (2015) tested this by having respondents rate the valence of positive (*e.g.,* I am very happy) and negative phrases (*e.g.,* I hate it) followed by varying numbers of exclamation marks. They found that the number of exclamation marks used in the phrase predicted the strength of positive or negative interpretations. Therefore, the comparatively high rate of exclamation marks in conjunction with heightened negative emotion words used by mass shooters may further emphasize their psychological and emotional immersion across writing contexts.

Cognitive Complexity and Processes

Based on the results from previous studies indicating that mass shooters were higher in cognitive complexity relative to other groups (*e.g.*, Baele, 2017), shooters were predicted to score higher on the summary category, Analytic. This was supported in comparisons between the entire shooter sample and the LIWC2015 expressive writing baseline. This is an interesting and relatively uncommon pattern in the computational linguistics literature, as Analytic tends to negatively associate with emotionality (Cohn et al., 2004). However, this pattern of high cognitive complexity did not generalize to other comparisons in the present study. In fact, mass shooters scored similarly to the LIWC2015 blog posts and consistently lower than the prisoner sample.

Several LIWC2015 categories are useful when predicting cognitive complexity and processes. These categories rely heavily on function words, such as prepositions and articles, that are used to organize sentences (Pennebaker et al., 2014). Conversely, to compute Analytic scores, functions words used to create a charismatic, narrative structure are negatively weighted (*e.g.*, personal pronouns, adverbs, negations; Pennebaker, 2011a). Thus, it is not altogether surprising that mass shooters scored lower on Analytic relative to prisoners and blog posts considering shooters used more emotion words and first-person singular pronouns. Likewise, across all analyses, mass shooters scored lower on correlates of Analytic. In the exploratory analyses with all shooters' writings and shooters' blog posts, prepositions (positively weighted in Analytic) were selected as important, negative predictors of mass shooters. Similarly, word count, an index of verbal fluency and communicative frequency (Tausczik & Pennebaker, 2010), was selected in the exploratory analyses. It is notable that semicolons, used to create more complex sentence structures, were selected as important predictors in the model using shooters' blog posts.

Cognitive complexity can also be measured using content categories such as cognitive processes in LIWC2015 (*i.e.*, certainty, causal). Certainty, for example, tends to indicate lower complexity (Pennebaker, 2011a) and is often associated with emotional stability (Tausczik & Pennebaker, 2010). As hypothesized, mass shooters used more certainty words than the LIWC2015 expressive writing and blog post baselines. However, there was no difference between shooters and prisoners in rates of certainty words, although prisoners did use more causal words compared to shooters' blog posts.

Together, how to interpret these results is not entirely clear. A viable and likely explanation is the different circumstances of the prisoner and shooter samples. For prisoners, they may be less emotionally attached to the topic discussed in their writings. This notion is supported by their lower emotionality and higher rates of Analytic and causal words. In text-analyses of violent extremist organizations across time, Pennebaker (2011a) found that complex thinking was lower prior to a terrorist attack and increased in the months following the attack. Since the individuals included in the prisoner sample had already committed violent acts, it is likely that the pattern found by Pennebaker (2011a) applies here as well. Therefore, these results may be a byproduct of the comparison sample chosen for the present study.

Pronouns can also be used to indicate attention allocation (Tausczik & Pennebaker, 2010). Relative to women, men tend to use more third-person plural pronouns when describing an event in which they were teased (Kowalski, 2000). According to Langman (2015) and others (*e.g.,* Neuman et al., 2015), mass shooters tend to have a heightened perception that they have been treated unfairly. Therefore, shooters were predicted to use more third-person plural pronouns relative to the other groups. This was the case for comparisons with the LIWC2015 baselines, yet there was no difference in third-person pronoun rates between prisoners and shooters. However, it is possible that this is a consequence of the comparison sample used in this study, as all of the individuals in the prisoner sample were men who potentially had a similar perspective. Nonetheless, it is notable that third-person singular pronouns (*e.g.,* he, she), a potential alternative to third-person plural in this context, were positive predictors of the shooter group.

Specific Personal Concerns

One of the criticisms levied against previous studies (*e.g.,* Langman, 2009a; Dutton et al., 2013) in the introduction was that they focus too heavily on predicting traits. Instead, an individual's written artifacts are likely to be better suited for examining the things with which they are most concerned (*i.e.,* personal concerns; McAdams, 1995; Chung & Pennebaker, 2018). In the present study, mass shooters were hypothesized to be highly concerned with power, as many shooters cite their foiled achievements as a source of outrage (Lankford, 2015). Furthermore, Kaati et al. (2016) found that both power and drives (broader category encompassing power and other motivations) were markedly high among this population. This was the case in analyses comparing the entire corpus of mass shooters to blog posts and expressive writing baselines. However, the differences in reference to power were removed when only comparing shooters journal entries and blogs to the LIWC2015 baselines. Thus, the writing-type coding procedure may have revealed some important distinctions between writing contexts for mass shooters.

Across all analyses, mass shooters also made references to power (and drives more broadly) less than the prisoner sample. As with some of the results discussed earlier, this is potentially a consequence of the comparison sample used here. Prisoners are in a position of relatively low power, and they are likely to be often reminded of their low standing in the social hierarchy. Therefore, the results for power may better represent a salient personal concern for prisoners rather than mass shooters. Still, it is interesting that shooters scored no differently than the expressive writing and blog post baselines once the other writing types were removed (manifestos, letters/emails, social media posts).

The exploratory analyses also revealed some potentially important personal concerns among shooters. These results indicated that mass shooters preferentially focused on friends and sex. The heightened use of references to friends contradicts previous work (Kaati et al., 2016), and it is unclear whether this focus on social affiliations is due to a perceived lack of friends (as has been described qualitatively in Langman, 2015 and quantitatively in McGee & DeBernardo, 1999). Newman and Fox (2009) also identified social marginalization as a potential risk factor for school shootings.

Likewise, Egnoto and Griffin (2016) hypothesized that sex would be a salient issue for shooters, although they found no differences between shooters and suicide notes in references to sex. Indeed, shooters are often cited as being victims of sexual abuse (Langman, 2009, 2009a, 2015), and sex is often discussed in narratives related to trauma, depression, suicidality, and violence (Tausczik & Pennebaker, 2010). This may be supported by some of the results of the present study. In exploratory analyses including all writing types, mass shooters used 53 times (*OR* = 53.10) more references to sex. However, this effect did not generalize to other exploratory analyses.

Finally, hypotheses for the Moral Foundations Dictionary (MFD; Graham et al., 2009) were largely unsupported. Morality, primarily harm and fairness, was expected to an especially important personal concern for mass shooters considering their tendency to rationalize their eventual attack (Lankford, 2013). However, shooters scored no different than prisoners for all of the MFD categories. Shooters may still express a distinct constellation of moral foundations in their writings, but the MFD was unable to detect it. In exploratory analyses including all shooters' writings, the MoralityGeneral category was selected as an important predictor. In fact, shooters made approximately 170 times more references to morality than prisoners. MoralityGeneral, as opposed to the other MFD categories, is larger and likely had more predictive power. This effect did not occur in other analyses, although it may still indicate morality is an important personal concern for mass shooters.

Limitations

There are a number of reasons to be cautious when interpreting the results of the present study. Foremost, the software used to analyze shooters' and prisoners' writings does not capture many important aspects of verbal and written communications. Namely, LIWC2015 does not measure context. This means that researchers must control writing context experimentally (*i.e.*, with specific writing prompts) or that the writing context must be publicly known. This was the case in the present study among the prisoner sample who all responded to the same, open-ended prompt. Likewise, the coding scheme used herein was an attempt to control this writing context. Still, the context of certain words tends to be lost within the LIWC2015 software, because words are not measured in relation to the other words co-occurring within the same text. For example, the word “fly” functions quite differently in the phrase, “fly away,” compared to “fly on the wall.” But, LIWC2015 does not distinguish between separate meanings of homonyms, rather it would count the word “fly” as matching multiple word categories. Thus, it is likely that alternative software options that account for such conversational variance would likely be useful in future studies. Similarly, the MFD predicted few differences between shooters and prisoners whatsoever. Indeed, the only MFD category identified as important was MoralityGeneral in exploratory analyses including the entire corpus of shooters’ writings, which is a likely consequence of the increased power for this category relative to the smaller, moral foundations. It is my view that morality is still an important construct to consider in studies on mass shooters. However, the MFD as it currently stands is insufficient due to its small size.

It is also worth noting that the analyses conducted in this study are mostly exploratory and should be treated as such. Results of exploratory analyses using stepwise logistic regression may not generalize to comparisons between shooters and other groups. There is also a high likelihood that multiple word-categories selected by the stepwise procedure were coincidentally significant, while actual explanatory variables were removed (Malek et al., 2007; Smith, 2018). Some individuals in the mass shooter sample were also overrepresented—that is, they wrote more across multiple writing types. Therefore, it is possible that the results of this study better reflect differences between prisoners and a subgroup of mass shooters. It is also important to consider that not all mass shooters had analyzable communications, and there may be distinct, unidentifiable differences between the present sample and shooters who were not included.

Lastly, it is unclear whether an appropriate comparison sample was used in this study. While offenders of violent crimes appear to be an appropriate control group, there are several issues with comparing prisoners' post-offense communications to mass shooters' pre-offense communications. First, the prisoner sample likely had time to reflect on their present situation in relation to a greater socio-cultural context. This is potentially reflected in their higher Analytic scores, which can be an index of psychological detachment from a discussion topic (Pennebaker et al., 2014). While a novel finding, as mass shooters typically display high cognitive complexity relative to comparison groups, it is notable that the low Analytic scores and higher negative emotion scores found in the shooter sample were potentially caused by differences between shooters' and prisoners' temporal relationship to their crimes. Future studies should consider experimental prompts for developing comparison groups, or, if available, using communications written by individuals prior to engaging in other forms of violent crime.

Conclusion

Despite these limitations, the present study has a number of strengths. This represents the largest and most selective sample of mass shooters to date. Compared to previous studies (*e.g.,* Baele, 2017; Egnoto & Griffin, 2016; Kaati et al., 2016), the data used herein were not confounded by the inclusion of individuals from other forms of mass violence (*e.g.,* serial murder). Considering that mass shooters are unique in terms of age, number of victims, base rate, and rate of death by suicide (Duwe, 2020), using strict exclusionary criteria is imperative when making inferences about the population. Coding for writing type was another major strength, because it allowed for alternative comparisons and conclusions. In previous studies, mass shooters communications have been compared to suicide notes (Egnoto & Griffin, 2016), civil rights activists (Baele, 2017), and the expressive writing baseline in LIWC2015. However, neither Egnoto and Griffin (2016) nor Baele (2017) controlled for the type of communication. Therefore, some blog posts were included in comparisons with the expressive writing baseline, whereas comparing them to the blog post baseline would have been more appropriate. By using the writing type coding scheme, a different pattern of results emerged that may help expand the limited nomological net surrounding mass shooters' personal concerns across contexts. Future linguistic analyses on mass shooters should therefore consider a similar coding procedure to develop hypotheses and identify appropriate comparison groups.

Overall, it is unlikely that the results of this study—or any other study—will lead to future prediction of mass shootings. However, as individuals continue to commit mass shootings, it will be important to further our understanding of mass shooters' psychological processes. And, computational linguistic analyses are one of the few windows through which we can begin to understand these constructs from the mass shooter's perspective.

References

Aharoni, E., Antonenko, O., & Kiehl, K. A. (2011) Disparities in the moral intuitions of criminal offenders: The role of psychopathy. *Journal of Research in Personality*, *45*(3), 322-327. https://doi.org/10.1016/j.jrp.2011.02.005

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Araque, O., Gatti, L., & Kalimeri, K. (2020). Moral strength: Exploiting a moral lexicon and embedding similarity for moral foundations prediction. *Knowledge-Based Systems, 191*, Article 105184. https://doi.org/10.1016/j.knosys.2019.105184

Baele, S. J. (2017). Lone-actor terrorists' emotions and cognition: An evaluation beyond stereotypes. *Political Psychology, 38*(3), 449-468. https://doi.org/10.1111/pops.12365

Behavioral Analysis Unit-2. (2008). *Serial murder: Multi-disciplinary perspectives for investigators*. Federal Bureau of Investigation. https://www.fbi.gov/stats-services/publications/serial-murder

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple hypothesis testing. Journal of the Royal Statistical Society, 57(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Carey, A. L., Brucks, M. S., Küfner, A. C. P., Holtzman, N. S., große Deters, F., Back, M. D., Donnellan, M. B., Pennebaker, J. W., & Mehl, M. R. (2015). Narcissism and the use of personal pronouns revisited. *Journal of Personality and Social Psychology, 109*(3), e1–e15. https://doi.org/10.1037/pspp0000029

Chung, C. K., and Pennebaker, J. W. (2008). Revealing dimensions of thinking in open-ended self-descriptions: an automated meaning extraction method for natural language. *Journal of Research in Personality, 42*(1), 96–132. https://doi.org/10.1016/j.jrp.2007.04.006

Dostoyevsky, F. (2008). *Demons*. Penguin Classics (Original work published 1871)

Dutton, D. G., White, K. R., Fogarty, D. (2013). Paranoid thinking in mass shooters. *Aggression and Violent Behavior*, *18*(5), 548-553. http://dx.doi.org/10.1016/j.avb.2013.07.012

Duwe, G. (2007). *Mass murder in the United States: A History*. McFarland & Company, Inc.

Duwe, G. (2020). Patterns of lethal mass violence. *Criminology & Public Policy*, *19*, 17-35. https://doi.org/10.1111/1745-9133.12478

Egnoto, M. J., & Griffin, D. J. (2016). Analyzing language in suicide notes and legacy tokens. *Crisis*, *37*(2), 140-147. https://doi.org/10.1027/0227-5910/a000363

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *PNAS, 115*(44), 11203-11208. https://doi.org/10.1073/pnas.1802331115

Federal Bureau of Investigation. (2010). *Crime in the United States.* https://ucr.fbi.gov/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/violent-crime

Glenn, A. L., Iyer, R., Graham, J., Koleva, S., & Haidt, J. (2009). Are all types of morality compromised in psychopathy? *Journal of Personality Disorders*, *23*(4), 384–398. https://doi.org/10.1521/pedi.2009.23.4.384

Goethe, J. W. v. (1985). *Faust*. Bantam Classics (Original work published 1808).

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029–46. https://doi.org/10.1037/a0015141

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., Ditto, P. H. (2013). Chapter two – Moral Foundations Theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55-130. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham, J., Nosek, B. A., & Haidt, J. (2012). The moral stereotypes of liberals and conservatives: Exaggeration of differences across the political spectrum. *PLoS One*, *7*, Article e50092. https://doi.org/10.1371/journal.pone.0050092

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98–116. https://doi.org/10.1007/s11211-007-0034-z

Hancock, J. T., Woodworth, M. T., Porter, S. (2013). Hungry like the wolf: A word pattern analysis of the language of psychopaths. *Legal and Criminological Psychology, 18*, 102-114. https://doi.org/10.1111/j.2044-8333.2011.02025.x

Handelman, L. D., & Lester, D. (2007). The content of suicide notes from attempters and completers. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, *28*(2), 102–104. https://doi.org/10.1027/0227-5910.28.2.102

Haynes, W. (2013). Benjamini-Hochberg method. In W. Dubitzky, O. Wolkenhauer, K. -H. Cho, & H. Yokota (Eds.), Encyclopedia of Systems Biology (2013 ed.). https://doi.org/10.1007/978-1-4419-9863-7_1215

Hirsh, J. B., & Peterson, J. B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality*, *43(*3), 524-527. https://doi.org/10.1016/j.jrp.2009.01.006

Holmes, D., Alpers, G. W., Ismailji, T., Classen, C., Wales, T., Cheasty, V., Miller, A., Koopman, C. (2007). Cognitive and emotional processing in narratives of women abused by intimate partners. *Violence Against Women*, *13*(11), 1192-1205. https://doi.org/10.1177/1077801207307801

Holtzman, N. S., Tackerman, A. M., Carey, A. L., Brucks, M. S., Kufner, A. C. P., Deters, F. G., Back, M. D., Donnellan, M. B. Pennebaker, J. W., Sherman, R. A., & Mehl, M. R. (2019). Linguistic markers of grandiose narcissism: A LIWC analysis of 15 samples. *Journal of Language and Social Psychology, 38*(5-6), 773-786. https://doi.org/10.1177/0261927X19871084

Ireland, M. E., & Mehl, M. R. (2014). Natural language use as a marker of personality. In T. Holtgraves (Ed.), *Oxford handbook of language and social psychology* (pp. 201–18). Oxford University Press.

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432. https://doi.org/10.1037/a0019265

Jonason, P. K., Strosser, G. L., Kroll, C. H., Duineveld, J. J., & Baruffi, S. A. (2015). Valuing myself over others: The Dark Triad traits and moral and social values. *Personality and Individual Differences, 81*, 102-106. https://doi.org/10.1016/j.paid.2014.10.045

Kaati, L., Shrestha, A., Cohen, K. (2016, June 12–June 14). *Linguistic analysis of lone offender manifestos* [Paper presentation]. IEEE 4th Annual International Conference on Cybercrime and Computer Foresics, Vancouver, BC. Canada.

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. Journal of Language and Social Psychology, 33(2), 125-143. https://doi.org/10.1177/0261927X13502654

Kelly, R. (2012). *Active shooter report: Recommendations and analysis for risk mitigation.* New York City Police Department. http://www.nyc.gov/html/nypd/downloads/pdf/counterterrorism/ActiveShooter2012Edition.pdf

Kennedy, B., Atari, M., Davani, A. M., Hoover, J., Omrani, A., Graham, J., Dehghani, M. (2021). Moral concerns of differentially observable in language. *Cognition*, *212*, Article 104696. https://doi.org/10.1016/j.cognition.2021.104696

Kim, K., Choi, S., Lee, J., & Sea, J. (2019). Differences in linguistic and psychological characteristics between suicide notes and diaries. *Journal of General Psychology*, *146*(4), 391-416. https://doi.org/10.1080/00221309.2019.1590304

Kop, M., Read, P., & Walker, B. R. (2019). Pseudocommando mass murderers: A big five personality profile using psycholinguistics. *Current Psychology*. https://doi.org/10.1007/s12144-019-00230-z

Kowalski, R. M. (2000). “I was only kidding!”: Victims’ and perpetrators’ perceptions of teasing. *Personality and Social Psychology Bulletin*, *26*(2), 231-241. https://doi.org/10.1177/0146167200264009

Kowalski, R. M. (2000). “I was only kidding!”: Victims’ and perpetrators’ perceptions of teasing. *Personality & Social Psychology Bulletin*, *26*(2), 231-241. https://doi.org/10.1177/0146167200264009

Langman, P. (2009a). Rampage school shooters: A typology. *Aggression and Violent Behavior. 14*(1), 79–86. https://doi.org/10.1016/j.avb.2008.10.003

Langman, P. (2009b). *Why kids kill: Inside the minds of school shooters.* St. Martin's Press.

Langman, P. (2015). *School shooters: Understanding high school, college, and adult perpetrators.* Rowman & Littlefield.

Lankford, A. (2013). A comparative analysis of suicide terrorists and rampage, workplace, and school shooters in the United States from 1990 to 2010. *Homicide Studies, 17*(3), 255-274. https://doi.org/10.1177/1088767912462033

Lankford, A. (2015). Mass shooters in the USA, 1966–2010: Differences between attackers who live and die. *Justice Quarterly, 32*(2), 360-379. https://doi.org/10.1080/07418825.2013.806675

Lindberg, N., Oksanen, A., Sailas, E., & Kaltiala-Heino, R. (2012). Adolescents expressing school massacre threats online: Something to be extremely worried about? *Child and Adolescent Psychiatry and Mental Health*, *6*(39). https://doi.org/10.1186/1753-2000-6-39

Malek, M. H., Berger, D. E., & Coburn, J. W. (2007). On the inappropriateness of stepwise regression analysis for model building and testing. *European Journal of Applied Physiology, 101*, 263-264. https://doi.org/10.1007/s00421-007-0485-9

Marshall, J., Watts, A. L., & Lilienfeld, S. O. (2018). Do psychopathic individuals possess a misaligned moral compass? A meta-analytic examination of psychopathy's relations with moral judgment. *Personality Disorders: Theory, Research, and Treatment, 9*(1), 40–50. https://doi.org/10.1037/per0000226

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality, 63*(3), 365-396. https://doi.org/10.1111/j.1467-6494.1995.tb00500.x

McAdams, D. P. (2006). *The redemptive self: Stories Americans live by*. Oxford University Press.

McGee, J. P., & DeBernardo, C. R. (1999). The classroom avenger: A behavioral profile of school based shootings. *The Forensic Examiner, 8*(5-6), 16–18.

Metzl, J. M., & Macleish, K. T. (2015). Mental illness, mass shootings, and the politics of American firearms. *American Journal of Public Health, 105*(2), 240-249. https://doi.org/10.2105/AJPH.2014.302242

Milton, J. (1968). *Paradise lost & paradise regained*. Signet Classic.

Neuman, Y., Assaf, D., Cohen, Y., & Knoll, J. L. (2015). Profiling school shooters: Automatic text-based analysis. *Frontiers in Psychiatry, 6*(86), 1-5. https://doi.org/10.3389/fpsyt.2015.00086

Newman, K. S., Fox, C., Harding, D. J., Mehta, J., & Roth, W. (2004). *Rampage: The Social Roots of School Shootings*. Basic Books.

Newman, K., & Fox, C. (2009). Repeat tragedy: Rampage shootings in American high school and college settings, 2002–2008. *American Behavioral Scientist, 52*(9), 1286–1308. https://doi.org/10.1177/0002764209332546

O'Neill, J. P., Miller, J. J., Waters, J. R. (2016). *Active shooter report: Recommendations and analysis for risk mitigation.* New York City Police Department. Retrieved from: https://www1.nyc.gov/assets/nypd/downloads/pdf/counterterrorism/active-shooter-analysis2016.pdf

Patrick, C. J. (2007). Antisocial personality disorder and psychopathy. In W. O'Donohue, K. A. Fowler, & S. O. Lilienfeld (Eds.), *Personality disorders: Toward the DSM-V* (pp. 109–166). Sage Publications, Inc. https://doi.org/10.4135/9781483328980.n6

Pennebaker, J. W. (2011a). Using computer analyses to identify language style and aggressive intent: The secret life of function words. *Dynamics of Asymmetric Conflict*, *4*(2), 92-102. https://doi.org/10.1080/17467586.2011.627932

Pennebaker, J. W. (2011b). *The secret life of pronouns: What our words say about us*. Bloomsbury.

Pennebaker, J. W., & Chung, C. K. (2007). Computerized text analysis of Al-Qaeda transcripts. In K. Krippendorff & M. Bock (Eds.), *A Content Analysis Reader* (pp. 453–469). Sage.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count (LIWC): LIWC2007.* Austin, TX: LIWC.net

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic inquiry and word count: LIWC 2015*. Austin, TX: Pennebaker Conglomerates. Retrieved from www.LIWC.net

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PLoS ONE, 9*(12), Article e115844. https://doi.org/10.1371/journal.pone.0115844

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah: Lawrence Erlbaum Associates.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Reviews of Psychology*, *54*, 547-577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Rayson, P. (2008). *Wmatrix: A web-based corpus processing environment.* Computing Department, Lancaster University. Retrieved from http://ucrel.lancs.ac.uk/wmatrix/

Sapru, M. (2019). Aggrieved, entitled, and hostile: A word-level analysis of the writings of three rampage shooters. *gnovis*, *19*(2), 2-16.

Sexton, J. B., & Helmreich, R. L. (2000). Analyzing cockpit communication: The links between language, performance, error, and workload. *Journal of Human Performance in Extreme Environments*, *5*(1), 63-68.

Smith, G. (2018). Step away from stepwise. *Journal of Big Data*, *5,* Article 32. https://doi.org/10.1186/s40537-018-0143-6

Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*, *14*(2), 287-305. https://doi.org/10.1177/1094428110369842

Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, *63*(4), 517-522.

Tamborini, R., Eden, A., Bowman, N. D., Grizzard, M., & Lachlan, K. A. (2012). The influence of morality subcultures on the acceptance and appeal of violence. *Journal of Communication*, *62*(1), 136–157. https://doi.org/10.1111/j.1460-2466.2011.01620.x

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24-54. https://doi.org/10.1177/0261927X09351676

The, P. L., Rayson, P., Pak, I., & Piao, S. (2015, December 11–13). *Sentiment analysis tools should take account of the number of exclamation marks!!!* [Paper Presentation]. iiWAS Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services, Brussels, Belgium.

Van den Nest, M., Till, B., & Niederkrotenthaler, T. (2018). Comparing indicators of suicidality among users in different types of nonprofessional suicide message boards. *Crisis*, *40*(2), 125-133. https://doi.org/10.1027/0227-5910/a000540

Whissell, C. M., & Dewson, M. J. (1986). A dictionary of affect in language: III. Analysis of two biblical and two secular passages. *Perceptual and Motor Skills*, *62*, 127– 133.

Winegard, B., & Deaner, R. O. (2010). The evolutionary significance of Red Sox nation: Sport fandom as a byproduct of coalitional psychology. *Evolutionary Psychology*, *8*(3), 432–446.

Ye, S., Yang, Q., Lan, T., Wang, Y., Zhu, Y., Zhu, B., Dong, Y., Kruger, F. (2021). Psychopathic traits predict moral judgments in five moral domains: The mediating effect of unpleasantness. *Legal and Criminological Psychology*, *26*(2), 176-195. https://doi.org/10.1111/lcrp.12189

Table 1

Racial Demographic Characteristics of Shooter and Prisoner Samples

Baseline Characteristic	*N*	%
Mass Shooters	36	--
White	21	60.56
Black/African American	3	8.45
Latinx	4	11.27
Native American	2	4.23
Asian	3	7.04
Multiracial	3	8.45
Prisoners	36	--
White	23	63.88
Black/African American	5	13.89
Latinx	4	11.11
Native American	1	2.78
Asian	1	2.78
Multiracial	2	5.56

Table 2

Means, Standard Deviations, and Frequency of Identically Zero Scores Across Samples for Hypothesized LIWC2015 and MFD Categories

Word Categories	***M* (*SD*)**	**Frequency of 0%**[a]
LIWC Categories		
Dictionary Words	82.80 (10.16)	0 (0%)
Analytic	53.44 (27.13)	0 (0%)
they	1.02 (1.67)	85 (37.00%)
Negative emotion	3.84 (3.05)	19 (8.26%)
certainty	1.95 (1.74)	39 (17.39%)
power	2.79 (2.48)	33 (14.35%)
MFD Categories		
Dictionary Words	2.18 (2.44)	42 (18.26%)
Harm	.85 (2.05)	77 (34.48%)
Fairness	.11 (.33)	163 (70.87%)
Loyalty	.28 (.55)	122 (53.04%)
Authority	.25 (.43)	122 (53.04%)
Purity	.15 (.61)	156 (67.83%)

Note. LIWC2015 = Linguistic Inquiry and Word Count 2015 Dictionary; MFD = Moral Foundations Dictionary.

[a] Frequency of 0% indicates the rate and percentage that a single text matched 0 words in the corresponding category.

Table 3

Results of Binomial Logistic Regression Hypothesis Testing of Differences Between Shooters and Prisoners

Word Category	**B**	***SE***	***OR***	**95% CI for *OR***	
				LL	UL
LIWC2015					
Analytic	-0.04	0.01	0.96***	0.94	0.98
Third-person plural	-0.11	0.10	0.90	0.73	1.10
Negative Emotion	0.05	0.09	1.05	0.87	1.26
Certainty	-0.19	0.17	0.83	0.60	1.15
Power	0.18	0.08	0.83*	0.72	0.97
Swear	1.01	0.47	2.753*	1.09	6.96
MFD					
Harm	-0.20	0.27	0.82	0.48	1.41
Fairness	-0.51	0.63	0.60	0.17	2.08
Loyalty	-0.19	0.38	0.83	0.39	1.75
Authority	-0.64	0.40	0.53	0.24	1.16
Purity	-0.64	0.94	0.53	0.08	3.32

Note. Results from the analyses presented in Table 3 include mass shooters' writings from all types. LIWC2015 = Linguistic Inquiry and Word Count 2015 Dictionary; MFD = Moral Foundations Dictionary; *SE* = standard enrror; *OR* = odds ratio; *CI* = confidence interval; *LL* = lower limit; *UL* = upper limit.

*p < .05. **p < .01. ***p < .001.

Table 4

Results of Binomial Logistic Regression Models Comparing Shooters' Journals and Blog Posts to Prisoners

Word Category	B	*SE*	*OR*	95% CI for *OR*	
				LL	UL
LIWC2015: Journal					
Analytic	-0.06	0.02	0.94***	0.91	0.97
Third-person plural	-0.55	0.27	0.58*	0.34	0.98
Negative Emotion	-0.19	0.20	1.21	0.81	1.80
Certainty	0.02	0.25	1.02	0.63	1.65
Power	-0.29	0.19	0.75	0.51	1.09
Swear	1.00	0.53	2.72	0.96	7.76
MFD: Journal					
Harm	0.02	0.17	1.02	0.74	1.40
Fairness	-0.31	0.57	0.73	0.24	2.22
Loyalty	-1.24	0.74	0.29	0.07	1.22
Authority	-2.18	0.68	0.11**	0.03	0.43
Purity	0.13	0.41	1.14	0.51	2.54
LIWC2015: Blog					
Analytic	-0.02	0.01	0.98	0.95	1.00
Third-person plural	-0.05	0.10	0.96	0.78	1.17

Negative Emotion	-0.03	0.10	0.97	0.80	1.19
Certainty	-0.16	0.19	0.85	0.59	1.23
Power	-0.19	0.09	0.82*	0.70	0.98
Swear	1.06	0.51	2.88*	1.06	7.83
MFD: Blog					
Harm	-0.20	0.27	0.82	0.48	1.41
Fairness	-0.51	0.63	0.60	0.17	2.08
Loyalty	-0.19	0.38	0.83	0.39	1.75
Authority	-0.64	0.40	0.53	0.24	1.16
Purity	-0.64	0.94	0.53	0.08	3.32

Note. In all models: mass shooters = 1 and prisoners = 0; positive beta estimates and odd ratios greater than 1.00 indicate mass shooters were positively predicted by the word category. LIWC2015 = Linguistic Inquiry and Word Count 2015 Dictionary; MFD = Moral Foundations Dictionary; *SE* = standard error; *OR* = odds ratio; CI = confidence interval; LL = lower limit; UL = upper limit.

*p < .05. **p < .01. ***p < .001.

Table 5

Exploratory Stepwise Binomial Logistic Regression Models of Differences Between Shooters and Prisoners Across Writing Types

Word Category	**B**	***SE***	***OR***	**95% CI for OR**	
				LL	UL
Model 1[a]					
Word count	-0.001	0.00	0.999**	0.999	1.000
First-person singular pronouns	0.75	0.21	2.13***	1.41	3.22
Third-person singular pronouns	1.43	0.39	4.19***	1.96	8.93
Prepositions	-0.91	0.25	0.4***	0.25	0.65
Conjunctions	-0.98	0.32	0.37**	0.20	0.70
Affect	1.07	0.33	2.92**	1.54	5.51
Sad	-2.99	1.14	0.05**	0.005	0.46
Friend	3.22	1.07	25.12**	3.07	205.47
Sexual	3.97	1.40	53.1**	3.43	821.19
Drives	-0.75	0.24	0.47**	0.30	0.76
Risk	-1.10	0.60	0.33	0.11	1.02
Exclamation mark	3.48	1.40	32.57*	2.17	487.77
Morality general	5.14	2.12	169.95**	2.66	10862.14
Model 2[b]					
Word count	-0.002	0.001	1.00*	0.99	1.00
Analytic	-0.09	0.04	0.91*	0.85	0.98

Clout	-0.05	0.02	0.95**	0.91	0.99
Power	-1.26	0.48	0.28**	0.11	0.72
Home	-3.80	1.28	0.02**	0.00	0.28
Exclamation mark	4.71	1.87	111.16*	2.83	4360.74
Model 3[c]					
Prepositions	-0.95	0.31	0.39**	0.21	0.72
Friend	2.79	1.72	16.33	0.56	478.48
Causal	-1.71	0.72	0.18*	0.04	0.74
Drives	-1.09	0.33	0.34**	0.18	0.64
Leisure	2.71	0.99	15.03**	2.17	104.10
Swear	1.84	0.76	6.27*	1.42	27.68
Semicolon	-6.47	2.72	0.002*	0.00	0.32
Exclamation mark	4.75	2.22	115.81*	1.50	8994.99

Note. In all models: mass shooters = 1 and prisoners = 0; positive beta estimates and odd ratios greater than 1.00 indicate mass shooters were positively predicted by the word category. LIWC2015 = Linguistic Inquiry and Word Count 2015 Dictionary; MFD = Moral Foundations Dictionary; SE = standard error; OR = odds ratio; CI = confidence interval; LL = lower limit; UL = upper limit.

[a] In Model 1, all shooters' writings were included in the analysis. [b] In Model 2, only shooters' journal entries were included in the analysis. [c] In Model 3, only shooters' blog posts were included in the analysis.

*p < .05. **p < .01. ***p < .001.

Table 6

Descriptive Statistics Across All LIWC2015 and MFD Categories

Word Category	Sample	N	Mean	SD	Min.	Max.
WC	**Prisoner**	36	1551.03	1414.33	142.00	5351.00
	Mass Shooter	194	628.85	1169.78	10.00	6186.00
	Manifesto	20	1078.25	1527.95	10.00	6186.00
	Blog	76	689.97	1376.32	12.00	6003.00
	Social Media	6	2036.67	2667.94	68.00	5570.00
	Journal	84	352.49	430.86	25.00	2531.00
	Letter	8	770.50	744.75	39.00	2318.00
Analytic	**Prisoner**	36	72.95	18.36	36.70	97.96
	Mass Shooter	194	49.82	26.99	1.00	98.87
	Manifesto	20	41.10	28.26	1.92	86.84
	Blog	76	62.63	26.10	2.89	98.87
	Social Media	6	49.18	30.72	5.18	72.06
	Journal	84	40.72	22.28	1.00	97.80
	Letter	8	46.06	32.38	7.62	96.61
Clout	**Prisoner**	36	59.76	23.85	17.88	98.71
	Mass Shooter	194	43.87	27.27	1.07	99.00
	Manifesto	20	43.23	23.77	4.46	96.64
	Blog	76	50.80	26.38	1.48	99.00

	Social Media	6	61.47	32.08	5.30	91.21
	Journal	84	34.95	25.67	1.07	99.00
	Letter	8	60.13	30.29	16.57	96.34
Authentic	**Prisoner**	36	43.32	29.21	6.06	94.79
	Mass Shooter	194	52.75	31.71	1.00	99.00
	Manifesto	20	55.73	28.70	1.00	98.24
	Blog	76	39.84	31.31	1.00	99.00
	Social Media	6	28.45	35.54	2.12	97.06
	Journal	84	65.53	27.33	6.20	99.00
	Letter	8	52.06	30.31	15.91	88.68
Tone	**Prisoner**	36	19.03	14.90	1.00	53.97
	Mass Shooter	194	30.11	33.06	1.00	99.00
	Manifesto	20	24.77	22.19	1.00	76.08
	Blog	76	33.66	35.79	1.00	99.00
	Social Media	6	14.51	14.92	1.00	38.54
	Journal	84	27.59	33.05	1.00	99.00
	Letter	8	47.95	33.26	2.58	97.19
WPS	**Prisoner**	36	18.18	5.10	9.11	39.82
	Mass Shooter	194	16.35	10.91	4.75	96.00
	Manifesto	20	17.01	13.83	8.19	74.20
	Blog	76	18.38	10.84	4.75	66.00
	Social Media	6	29.98	32.81	8.89	96.00

	Journal	84	13.16	5.34	5.50	44.00
	Letter	8	18.84	7.83	7.26	27.93
Sixltr	**Prisoner**	36	19.13	6.39	7.04	31.00
	Mass Shooter	194	15.00	5.74	0.00	47.06
	Manifesto	20	13.88	4.63	7.95	23.29
	Blog	76	14.94	6.35	5.26	47.06
	Social Media	6	19.32	6.28	11.74	24.79
	Journal	84	14.66	4.63	0.00	27.68
	Letter	8	18.74	10.04	5.80	34.78
Dic	**Prisoner**	36	83.37	5.63	65.87	93.17
	Mass Shooter	194	82.69	10.80	26.32	97.92
	Manifesto	20	87.17	3.29	80.68	91.90
	Blog	76	75.88	12.97	26.32	95.00
	Social Media	6	86.63	8.79	75.79	97.92
	Journal	84	87.10	6.27	70.00	97.86
	Letter	8	86.90	5.83	74.78	93.17
function	**Prisoner**	36	51.30	5.16	35.70	59.72
	Mass Shooter	194	49.05	11.69	5.26	65.96
	Manifesto	20	55.13	4.88	45.99	65.76
	Blog	76	40.86	13.40	5.26	61.67
	Social Media	6	54.88	6.95	42.30	61.76
	Journal	84	53.99	6.44	35.19	65.96

	Letter	8	55.48	7.25	40.65	63.98
pronoun	**Prisoner**	36	12.62	4.18	5.47	20.65
	Mass Shooter	194	16.09	6.01	0.00	34.45
	Manifesto	20	18.49	4.91	10.92	30.00
	Blog	76	12.75	5.88	0.00	33.33
	Social Media	6	15.67	6.92	10.64	29.41
	Journal	84	18.55	4.97	4.88	34.45
	Letter	8	16.30	5.01	5.65	21.12
ppron	**Prisoner**	36	7.63	3.77	1.52	13.80
	Mass Shooter	194	10.85	5.31	0.00	31.93
	Manifesto	20	13.16	5.27	6.61	30.00
	Blog	76	8.56	5.10	0.00	26.32
	Social Media	6	10.14	5.70	6.25	20.59
	Journal	84	12.42	4.74	0.00	31.93
	Letter	8	10.81	5.59	2.93	20.51
i	**Prisoner**	36	3.13	3.39	0.00	10.40
	Mass Shooter	194	6.75	4.63	0.00	20.32
	Manifesto	20	8.26	3.15	2.74	15.73
	Blog	76	4.60	4.20	0.00	16.67
	Social Media	6	4.75	6.64	0.00	17.65
	Journal	84	8.51	4.40	0.00	20.32
	Letter	8	6.48	4.06	0.87	13.04

we	**Prisoner**	36	0.74	0.99	0.00	4.37
	Mass Shooter	194	0.46	0.87	0.00	4.97
	Manifesto	20	0.34	0.65	0.00	2.85
	Blog	76	0.26	0.81	0.00	4.97
	Social Media	6	0.57	0.75	0.00	2.08
	Journal	84	0.65	0.98	0.00	4.55
	Letter	8	0.45	0.52	0.00	1.45
you	**Prisoner**	36	1.79	2.65	0.00	10.53
	Mass Shooter	194	1.50	2.62	0.00	16.67
	Manifesto	20	2.17	3.05	0.00	10.00
	Blog	76	1.48	2.86	0.00	16.67
	Social Media	6	1.66	1.19	0.36	3.12
	Journal	84	1.22	2.22	0.00	15.00
	Letter	8	2.83	3.62	0.24	10.26
shehe	**Prisoner**	36	0.77	0.82	0.00	3.32
	Mass Shooter	194	1.16	2.06	0.00	13.45
	Manifesto	20	1.61	2.60	0.00	10.00
	Blog	76	1.14	2.06	0.00	9.31
	Social Media	6	1.48	2.36	0.00	6.11
	Journal	84	1.14	2.00	0.00	13.45
	Letter	8	0.21	0.27	0.00	0.72
they	**Prisoner**	36	1.20	0.99	0.00	3.90

	Mass Shooter	194	0.98	1.77	0.00	18.52
	Manifesto	20	0.79	0.75	0.00	2.66
	Blog	76	1.08	2.49	0.00	18.52
	Social Media	6	1.67	1.34	0.00	2.98
	Journal	84	0.91	1.14	0.00	4.50
	Letter	8	0.84	0.96	0.00	2.59
ipron	**Prisoner**	36	4.99	1.74	0.70	8.76
	Mass Shooter	194	5.23	2.96	0.00	17.14
	Manifesto	20	5.27	2.28	0.00	9.05
	Blog	76	4.19	3.06	0.00	16.67
	Social Media	6	5.53	2.71	1.96	8.82
	Journal	84	6.11	2.80	0.00	17.14
	Letter	8	5.48	2.79	0.00	8.84
article	**Prisoner**	36	7.76	1.73	2.61	11.48
	Mass Shooter	194	5.46	2.70	0.00	14.63
	Manifesto	20	5.35	2.32	0.00	8.40
	Blog	76	5.28	3.19	0.00	13.43
	Social Media	6	6.50	1.83	3.12	8.31
	Journal	84	5.55	2.32	1.23	14.63
	Letter	8	5.72	3.06	0.00	9.02
prep	**Prisoner**	36	14.16	1.47	10.53	18.20
	Mass Shooter	194	10.48	3.81	0.00	19.26

	Manifesto	20	11.36	2.61	7.14	16.04
	Blog	76	9.11	4.84	0.00	19.26
	Social Media	6	11.30	3.68	6.25	15.91
	Journal	84	11.24	2.55	3.57	19.23
	Letter	8	12.78	2.79	6.83	14.63
auxverb	**Prisoner**	36	7.79	1.88	4.44	11.28
	Mass Shooter	194	9.20	3.47	0.00	21.77
	Manifesto	20	10.20	3.66	2.92	20.00
	Blog	76	7.93	3.83	0.00	21.77
	Social Media	6	9.57	2.49	6.58	13.54
	Journal	84	9.84	2.64	3.70	16.00
	Letter	8	11.73	4.38	5.65	17.95
adverb	**Prisoner**	36	4.16	1.39	1.50	7.51
	Mass Shooter	194	4.60	2.71	0.00	13.87
	Manifesto	20	4.82	2.08	0.00	8.70
	Blog	76	3.57	2.85	0.00	11.32
	Social Media	6	5.88	3.84	3.64	13.54
	Journal	84	5.36	2.42	0.00	13.87
	Letter	8	5.06	2.09	2.07	7.97
conj	**Prisoner**	36	6.40	1.46	3.36	10.39
	Mass Shooter	194	5.23	3.19	0.00	15.38
	Manifesto	20	6.83	2.26	3.05	10.69

	Blog	76	4.14	3.37	0.00	12.96
	Social Media	6	7.34	3.12	5.44	13.54
	Journal	84	5.59	2.93	0.00	15.38
	Letter	8	6.17	3.04	2.56	13.04
negate	**Prisoner**	36	1.72	0.87	0.49	4.23
	Mass Shooter	194	1.92	1.76	0.00	11.11
	Manifesto	20	1.90	1.10	0.00	3.81
	Blog	76	1.45	1.87	0.00	11.11
	Social Media	6	1.88	0.74	1.04	3.18
	Journal	84	2.37	1.80	0.00	8.00
	Letter	8	1.77	0.73	0.22	2.56
verb	**Prisoner**	36	14.07	3.70	7.31	20.85
	Mass Shooter	194	16.31	4.45	2.78	27.43
	Manifesto	20	16.90	3.24	11.50	25.24
	Blog	76	14.32	4.52	2.78	24.53
	Social Media	6	15.54	3.69	11.76	22.06
	Journal	84	17.89	3.97	8.57	27.43
	Letter	8	17.78	5.07	10.76	26.81
adj	**Prisoner**	36	4.09	1.14	1.75	7.12
	Mass Shooter	194	4.09	2.29	0.00	11.48
	Manifesto	20	3.71	1.56	0.00	6.38
	Blog	76	3.82	2.50	0.00	9.76

	Social Media	6	4.02	1.26	2.20	5.88
	Journal	84	4.43	2.32	0.00	11.48
	Letter	8	4.13	1.97	0.72	7.69
compare	**Prisoner**	36	2.07	0.86	0.52	4.06
	Mass Shooter	194	1.90	1.42	0.00	9.84
	Manifesto	20	1.80	0.97	0.00	4.26
	Blog	76	1.66	1.53	0.00	5.26
	Social Media	6	1.62	1.00	0.00	2.76
	Journal	84	2.13	1.46	0.00	9.84
	Letter	8	2.20	0.71	0.72	2.88
interrog	**Prisoner**	36	1.80	0.97	0.00	4.66
	Mass Shooter	194	1.30	1.14	0.00	5.00
	Manifesto	20	1.41	0.75	0.00	3.23
	Blog	76	0.94	1.10	0.00	5.00
	Social Media	6	1.46	0.87	0.00	2.42
	Journal	84	1.58	1.22	0.00	4.72
	Letter	8	1.50	1.02	0.00	2.90
number	**Prisoner**	36	2.00	1.56	0.00	9.60
	Mass Shooter	194	4.73	5.13	0.00	27.03
	Manifesto	20	1.90	1.36	0.00	6.20
	Blog	76	7.40	6.50	0.00	27.03
	Social Media	6	0.84	0.94	0.00	2.52

	Journal	84	3.61	3.23	0.00	16.67
	Letter	8	1.28	1.28	0.00	2.90
quant	**Prisoner**	36	2.06	0.80	0.33	3.52
	Mass Shooter	194	1.97	1.54	0.00	8.33
	Manifesto	20	1.94	0.97	0.00	4.26
	Blog	76	1.69	1.80	0.00	8.33
	Social Media	6	2.10	1.39	0.00	4.17
	Journal	84	2.19	1.43	0.00	6.56
	Letter	8	2.25	0.81	0.72	3.35
affect	**Prisoner**	36	5.22	1.79	2.21	9.43
	Mass Shooter	194	8.08	4.74	0.00	28.79
	Manifesto	20	6.94	2.46	2.97	11.69
	Blog	76	9.05	6.06	0.00	28.79
	Social Media	6	9.48	6.92	4.79	22.92
	Journal	84	7.51	3.40	0.00	17.80
	Letter	8	6.61	4.07	2.17	15.38
posemo	**Prisoner**	36	2.12	1.09	0.49	5.35
	Mass Shooter	194	3.47	3.77	0.00	27.27
	Manifesto	20	2.96	1.70	0.00	7.06
	Blog	76	3.94	5.25	0.00	27.27
	Social Media	6	1.81	1.47	0.00	3.50
	Journal	84	3.26	2.46	0.00	15.13

	Letter	8	3.84	2.96	1.63	10.26
negemo	**Prisoner**	36	3.05	1.28	1.03	6.55
	Mass Shooter	194	3.99	3.25	0.00	21.88
	Manifesto	20	3.96	2.08	1.68	10.00
	Blog	76	3.63	3.54	0.00	16.67
	Social Media	6	7.46	7.68	1.96	21.88
	Journal	84	4.20	2.71	0.00	15.38
	Letter	8	2.68	1.59	0.00	5.13
anx	**Prisoner**	36	0.34	0.33	0.00	1.50
	Mass Shooter	194	0.35	0.90	0.00	9.09
	Manifesto	20	0.28	0.42	0.00	1.63
	Blog	76	0.39	1.21	0.00	9.09
	Social Media	6	0.23	0.20	0.00	0.50
	Journal	84	0.35	0.71	0.00	4.11
	Letter	8	0.32	0.31	0.00	0.84
anger	**Prisoner**	36	1.28	0.84	0.00	3.92
	Mass Shooter	194	2.31	2.80	0.00	16.67
	Manifesto	20	2.03	2.30	0.00	10.00
	Blog	76	2.15	3.34	0.00	16.67
	Social Media	6	4.50	6.14	0.79	16.67
	Journal	84	2.49	2.00	0.00	8.18
	Letter	8	0.99	1.05	0.00	2.83

sad	**Prisoner**	36	0.59	0.47	0.00	1.76
	Mass Shooter	194	0.56	0.80	0.00	3.85
	Manifesto	20	0.68	0.52	0.00	1.90
	Blog	76	0.33	0.64	0.00	2.78
	Social Media	6	1.28	1.37	0.13	3.12
	Journal	84	0.67	0.88	0.00	3.85
	Letter	8	0.69	0.83	0.00	2.56
social	**Prisoner**	36	8.96	3.63	3.89	16.32
	Mass Shooter	194	9.10	4.99	0.00	27.16
	Manifesto	20	10.25	4.23	0.00	20.00
	Blog	76	8.25	5.15	0.00	27.16
	Social Media	6	11.66	2.61	8.40	14.88
	Journal	84	9.13	4.90	0.00	25.21
	Letter	8	12.13	6.06	2.28	23.08
family	**Prisoner**	36	0.34	0.76	0.00	4.10
	Mass Shooter	194	0.42	0.77	0.00	5.13
	Manifesto	20	0.65	0.86	0.00	2.52
	Blog	76	0.15	0.32	0.00	1.23
	Social Media	6	0.67	0.52	0.14	1.47
	Journal	84	0.47	0.73	0.00	3.21
	Letter	8	1.72	1.99	0.00	5.13
friend	**Prisoner**	36	0.21	0.23	0.00	0.99

	Mass Shooter	194	0.41	0.76	0.00	5.71
	Manifesto	20	0.30	0.32	0.00	0.81
	Blog	76	0.25	0.53	0.00	2.97
	Social Media	6	0.33	0.57	0.00	1.47
	Journal	84	0.54	0.95	0.00	5.71
	Letter	8	0.85	0.97	0.00	2.56
female	**Prisoner**	36	0.36	0.64	0.00	3.32
	Mass Shooter	194	0.70	1.46	0.00	13.45
	Manifesto	20	1.00	1.49	0.00	6.52
	Blog	76	0.38	0.84	0.00	4.40
	Social Media	6	0.79	0.62	0.00	1.47
	Journal	84	0.90	1.89	0.00	13.45
	Letter	8	0.84	0.97	0.00	2.56
male	**Prisoner**	36	1.08	1.08	0.00	5.35
	Mass Shooter	194	1.06	1.72	0.00	10.00
	Manifesto	20	1.46	2.25	0.00	10.00
	Blog	76	1.30	2.04	0.00	9.16
	Social Media	6	1.54	1.72	0.00	4.89
	Journal	84	0.75	1.23	0.00	6.94
	Letter	8	0.64	0.87	0.00	2.56
cogproc	**Prisoner**	36	11.36	2.78	4.28	16.25
	Mass Shooter	194	11.02	5.39	0.00	23.17

	Manifesto	20	11.29	2.79	6.29	16.19
	Blog	76	8.26	5.58	0.00	23.17
	Social Media	6	11.01	4.68	3.12	15.46
	Journal	84	13.43	4.68	0.00	21.30
	Letter	8	11.31	4.11	5.13	17.60
insight	**Prisoner**	36	2.16	0.99	0.49	5.01
	Mass Shooter	194	2.14	1.65	0.00	6.67
	Manifesto	20	1.82	0.87	0.00	3.26
	Blog	76	1.42	1.62	0.00	5.66
	Social Media	6	2.15	1.03	1.04	3.26
	Journal	84	2.85	1.59	0.00	6.67
	Letter	8	2.37	1.43	0.00	4.35
cause	**Prisoner**	36	2.06	0.87	0.33	4.05
	Mass Shooter	194	1.36	1.28	0.00	6.52
	Manifesto	20	1.46	0.95	0.00	3.99
	Blog	76	1.02	1.36	0.00	6.52
	Social Media	6	1.85	1.54	0.00	4.41
	Journal	84	1.60	1.19	0.00	4.07
	Letter	8	1.56	1.33	0.00	3.11
discrep	**Prisoner**	36	1.38	0.70	0.25	2.91
	Mass Shooter	194	1.73	1.61	0.00	10.00
	Manifesto	20	2.20	2.03	0.00	10.00

	Blog	76	1.23	1.72	0.00	8.33
	Social Media	6	1.63	1.04	0.00	3.09
	Journal	84	2.12	1.33	0.00	5.95
	Letter	8	1.39	1.29	0.00	3.73
tentat	**Prisoner**	36	2.66	1.09	0.99	6.82
	Mass Shooter	194	2.37	1.78	0.00	10.00
	Manifesto	20	2.58	1.89	1.32	10.00
	Blog	76	1.56	1.71	0.00	7.52
	Social Media	6	1.79	1.48	0.00	3.69
	Journal	84	3.04	1.61	0.00	6.38
	Letter	8	3.00	0.85	1.45	4.40
certain	**Prisoner**	36	1.67	0.62	0.33	2.84
	Mass Shooter	194	2.00	1.87	0.00	11.11
	Manifesto	20	1.97	1.01	0.00	3.81
	Blog	76	1.89	2.30	0.00	11.11
	Social Media	6	1.56	0.81	0.00	2.20
	Journal	84	2.22	1.71	0.00	10.85
	Letter	8	1.23	0.77	0.00	2.37
differ	**Prisoner**	36	3.16	1.08	0.99	6.15
	Mass Shooter	194	2.84	2.17	0.00	12.24
	Manifesto	20	3.92	1.96	1.26	10.00
	Blog	76	1.99	2.30	0.00	12.24

	Social Media	6	3.56	1.01	2.08	4.70
	Journal	84	3.28	1.99	0.00	7.69
	Letter	8	3.18	1.02	2.17	4.97
percept	**Prisoner**	36	2.27	1.53	0.48	8.71
	Mass Shooter	194	2.93	2.52	0.00	15.00
	Manifesto	20	1.44	1.01	0.00	3.52
	Blog	76	4.43	2.99	0.00	15.00
	Social Media	6	1.60	0.99	0.00	3.12
	Journal	84	2.09	1.64	0.00	11.54
	Letter	8	2.19	1.63	0.62	5.88
see	**Prisoner**	36	0.78	0.66	0.00	3.27
	Mass Shooter	194	0.92	1.12	0.00	5.97
	Manifesto	20	0.48	0.64	0.00	2.64
	Blog	76	1.22	1.41	0.00	5.97
	Social Media	6	0.37	0.42	0.00	0.98
	Journal	84	0.82	0.84	0.00	3.42
	Letter	8	0.77	1.28	0.00	3.84
hear	**Prisoner**	36	0.76	1.44	0.00	8.71
	Mass Shooter	194	1.22	1.74	0.00	9.09
	Manifesto	20	0.52	0.50	0.00	1.43
	Blog	76	2.43	2.22	0.00	9.09
	Social Media	6	0.51	0.36	0.00	1.04

	Journal	84	0.42	0.55	0.00	2.26
	Letter	8	0.56	0.34	0.00	0.98
feel	**Prisoner**	36	0.51	0.54	0.00	2.42
	Mass Shooter	194	0.69	1.31	0.00	11.54
	Manifesto	20	0.39	0.46	0.00	1.43
	Blog	76	0.68	1.43	0.00	10.00
	Social Media	6	0.67	0.88	0.00	2.08
	Journal	84	0.77	1.41	0.00	11.54
	Letter	8	0.63	0.84	0.00	2.56
bio	**Prisoner**	36	2.28	1.32	0.52	6.34
	Mass Shooter	194	3.35	2.84	0.00	20.83
	Manifesto	20	2.68	1.66	0.00	6.38
	Blog	76	3.68	3.68	0.00	20.83
	Social Media	6	3.12	1.44	1.68	5.88
	Journal	84	3.36	2.30	0.00	11.54
	Letter	8	1.87	0.75	0.97	3.02
body	**Prisoner**	36	0.62	0.57	0.00	2.42
	Mass Shooter	194	1.02	1.57	0.00	13.33
	Manifesto	20	0.76	0.93	0.00	4.26
	Blog	76	1.28	2.04	0.00	13.33
	Social Media	6	0.62	0.76	0.00	1.71
	Journal	84	0.94	1.26	0.00	7.69

	Letter	8	0.30	0.38	0.00	1.08
health	**Prisoner**	36	1.10	0.65	0.00	2.69
	Mass Shooter	194	0.80	1.13	0.00	8.33
	Manifesto	20	1.08	0.59	0.00	2.25
	Blog	76	0.85	1.40	0.00	8.33
	Social Media	6	1.30	0.96	0.52	3.12
	Journal	84	0.68	0.99	0.00	5.56
	Letter	8	0.61	0.54	0.00	1.45
sexual	**Prisoner**	36	0.07	0.13	0.00	0.61
	Mass Shooter	194	0.75	1.81	0.00	16.67
	Manifesto	20	0.38	0.60	0.00	2.54
	Blog	76	0.84	2.65	0.00	16.67
	Social Media	6	1.12	1.16	0.00	2.94
	Journal	84	0.76	1.00	0.00	3.64
	Letter	8	0.28	0.79	0.00	2.24
ingest	**Prisoner**	36	0.47	0.75	0.00	3.52
	Mass Shooter	194	0.39	0.96	0.00	9.09
	Manifesto	20	0.15	0.26	0.00	1.12
	Blog	76	0.47	1.20	0.00	9.09
	Social Media	6	0.06	0.10	0.00	0.26
	Journal	84	0.43	0.88	0.00	6.19
	Letter	8	0.07	0.15	0.00	0.44

drives	**Prisoner**	36	8.24	1.91	4.75	13.88
	Mass Shooter	194	6.42	3.30	0.00	21.09
	Manifesto	20	7.62	3.46	3.97	20.00
	Blog	76	5.28	3.48	0.00	21.09
	Social Media	6	8.49	4.60	3.91	16.67
	Journal	84	6.86	2.74	1.63	20.00
	Letter	8	7.89	2.73	3.96	12.82
affiliation	**Prisoner**	36	1.61	1.14	0.27	5.53
	Mass Shooter	194	1.65	1.81	0.00	7.69
	Manifesto	20	1.55	1.41	0.00	4.74
	Blog	76	1.06	1.58	0.00	7.23
	Social Media	6	1.63	0.53	0.73	2.08
	Journal	84	2.05	1.90	0.00	7.56
	Letter	8	3.30	2.39	0.24	7.69
achieve	**Prisoner**	36	1.43	0.76	0.00	3.26
	Mass Shooter	194	0.89	0.96	0.00	6.67
	Manifesto	20	0.97	0.80	0.00	2.66
	Blog	76	0.58	0.74	0.00	2.97
	Social Media	6	0.77	0.60	0.00	1.22
	Journal	84	1.14	1.13	0.00	6.67
	Letter	8	1.25	0.77	0.00	2.56
power	**Prisoner**	36	4.04	1.35	1.77	7.24

	Mass Shooter	194	2.56	2.58	0.00	21.09
	Manifesto	20	3.16	2.06	0.00	10.00
	Blog	76	2.50	2.98	0.00	21.09
	Social Media	6	5.48	4.78	1.47	14.58
	Journal	84	2.30	2.04	0.00	16.67
	Letter	8	2.26	1.40	0.62	4.53
reward	**Prisoner**	36	1.10	0.67	0.00	2.63
	Mass Shooter	194	1.15	1.09	0.00	5.00
	Manifesto	20	1.32	0.90	0.00	2.92
	Blog	76	0.95	1.22	0.00	5.00
	Social Media	6	0.53	0.54	0.00	1.22
	Journal	84	1.33	1.02	0.00	4.88
	Letter	8	1.31	0.92	0.00	3.11
risk	**Prisoner**	36	0.84	0.48	0.29	2.11
	Mass Shooter	194	0.53	0.98	0.00	10.00
	Manifesto	20	1.14	2.15	0.00	10.00
	Blog	76	0.47	0.96	0.00	5.00
	Social Media	6	0.62	0.52	0.00	1.47
	Journal	84	0.44	0.49	0.00	2.01
	Letter	8	0.40	0.36	0.00	1.13
focuspast	**Prisoner**	36	3.21	2.61	0.00	10.15
	Mass Shooter	194	3.38	2.97	0.00	15.38

	Manifesto	20	3.90	2.35	0.00	10.00
	Blog	76	3.13	3.00	0.00	10.74
	Social Media	6	5.63	4.29	2.10	11.76
	Journal	84	3.40	3.02	0.00	15.38
	Letter	8	2.51	1.83	0.00	5.20
focuspresent	**Prisoner**	36	8.98	3.32	3.25	17.81
	Mass Shooter	194	10.81	4.16	0.00	21.77
	Manifesto	20	10.86	4.23	0.00	19.15
	Blog	76	9.35	4.25	0.00	21.77
	Social Media	6	8.36	2.15	5.21	10.29
	Journal	84	12.14	3.69	3.61	21.13
	Letter	8	12.49	4.29	7.07	19.57
focusfuture	**Prisoner**	36	1.11	0.75	0.14	3.99
	Mass Shooter	194	1.56	1.52	0.00	8.82
	Manifesto	20	1.85	1.37	0.00	4.49
	Blog	76	1.06	1.27	0.00	5.26
	Social Media	6	1.41	0.79	0.84	2.94
	Journal	84	1.86	1.56	0.00	8.82
	Letter	8	2.50	2.66	0.72	8.70
relativ	**Prisoner**	36	13.65	3.45	8.12	20.51
	Mass Shooter	194	12.67	5.20	0.00	36.11
	Manifesto	20	11.73	4.51	0.00	21.53

	Blog	76	12.83	5.93	1.36	36.11
	Social Media	6	9.50	1.91	7.35	12.50
	Journal	84	13.15	4.80	3.21	28.57
	Letter	8	11.01	4.42	5.13	18.12
motion	**Prisoner**	36	1.92	0.91	0.00	4.95
	Mass Shooter	194	1.47	1.37	0.00	6.67
	Manifesto	20	1.71	1.02	0.00	3.37
	Blog	76	1.50	1.61	0.00	6.67
	Social Media	6	1.70	2.11	0.00	5.88
	Journal	84	1.37	1.20	0.00	5.00
	Letter	8	1.39	0.77	0.60	2.56
space	**Prisoner**	36	7.34	2.13	4.19	15.67
	Mass Shooter	194	5.48	3.02	0.00	25.00
	Manifesto	20	5.07	1.79	0.00	8.51
	Blog	76	5.88	4.08	0.00	25.00
	Social Media	6	4.53	1.87	1.47	7.33
	Journal	84	5.33	2.17	0.00	11.43
	Letter	8	4.95	1.86	2.56	7.79
time	**Prisoner**	36	4.57	1.62	2.00	9.05
	Mass Shooter	194	5.81	3.66	0.00	25.71
	Manifesto	20	5.14	2.72	0.00	12.04
	Blog	76	5.48	3.12	0.00	16.67

	Social Media	6	3.43	2.76	0.00	8.33
	Journal	84	6.53	4.22	0.00	25.71
	Letter	8	4.85	3.68	0.00	11.59
work	**Prisoner**	36	2.13	1.21	0.38	4.90
	Mass Shooter	194	1.54	1.86	0.00	10.78
	Manifesto	20	2.34	2.32	0.00	8.55
	Blog	76	1.06	1.86	0.00	10.78
	Social Media	6	1.33	1.08	0.00	2.94
	Journal	84	1.80	1.73	0.00	10.00
	Letter	8	1.37	1.42	0.00	3.59
leisure	**Prisoner**	36	0.61	0.51	0.00	2.65
	Mass Shooter	194	1.72	2.41	0.00	18.18
	Manifesto	20	0.57	0.70	0.00	2.21
	Blog	76	2.97	3.17	0.00	18.18
	Social Media	6	0.76	0.83	0.00	2.10
	Journal	84	1.01	1.33	0.00	7.08
	Letter	8	0.88	0.72	0.00	2.48
home	**Prisoner**	36	0.53	0.52	0.00	2.30
	Mass Shooter	194	0.30	0.63	0.00	5.56
	Manifesto	20	0.33	0.43	0.00	1.28
	Blog	76	0.35	0.83	0.00	5.56
	Social Media	6	0.45	0.54	0.00	1.47

	Journal	84	0.18	0.39	0.00	2.27
	Letter	8	0.78	0.71	0.00	1.86
money	**Prisoner**	36	0.51	0.54	0.00	1.99
	Mass Shooter	194	0.37	0.69	0.00	4.35
	Manifesto	20	0.43	0.42	0.00	1.11
	Blog	76	0.31	0.76	0.00	4.35
	Social Media	6	0.23	0.33	0.00	0.84
	Journal	84	0.42	0.70	0.00	3.57
	Letter	8	0.34	0.50	0.00	1.41
relig	**Prisoner**	36	0.39	0.53	0.00	2.11
	Mass Shooter	194	0.67	1.91	0.00	21.09
	Manifesto	20	1.03	2.29	0.00	10.00
	Blog	76	0.83	2.70	0.00	21.09
	Social Media	6	0.32	0.38	0.00	0.98
	Journal	84	0.51	0.74	0.00	3.25
	Letter	8	0.23	0.35	0.00	0.97
death	**Prisoner**	36	0.29	0.51	0.00	2.97
	Mass Shooter	194	0.87	1.70	0.00	15.62
	Manifesto	20	1.44	2.18	0.00	10.00
	Blog	76	0.83	1.47	0.00	7.89
	Social Media	6	3.22	6.14	0.00	15.62
	Journal	84	0.66	0.91	0.00	4.88

	Letter	8	0.33	0.26	0.00	0.65
informal	**Prisoner**	36	0.44	0.50	0.00	2.30
	Mass Shooter	194	2.59	3.77	0.00	27.27
	Manifesto	20	0.99	1.29	0.00	5.32
	Blog	76	3.65	5.15	0.00	27.27
	Social Media	6	1.96	4.09	0.00	10.29
	Journal	84	2.26	2.31	0.00	11.38
	Letter	8	0.37	0.43	0.00	1.24
swear	**Prisoner**	36	0.15	0.32	0.00	1.41
	Mass Shooter	194	1.32	2.26	0.00	16.67
	Manifesto	20	0.59	1.23	0.00	5.32
	Blog	76	1.49	2.95	0.00	16.67
	Social Media	6	0.80	1.77	0.00	4.41
	Journal	84	1.48	1.77	0.00	7.41
	Letter	8	0.10	0.22	0.00	0.62
netspeak	**Prisoner**	36	0.11	0.21	0.00	0.89
	Mass Shooter	194	0.83	2.81	0.00	24.24
	Manifesto	20	0.13	0.22	0.00	0.63
	Blog	76	1.70	4.29	0.00	24.24
	Social Media	6	1.04	2.38	0.00	5.88
	Journal	84	0.27	0.49	0.00	2.27
	Letter	8	0.15	0.26	0.00	0.72

assent	**Prisoner**	36	0.08	0.13	0.00	0.50
	Mass Shooter	194	0.17	0.59	0.00	5.26
	Manifesto	20	0.06	0.10	0.00	0.30
	Blog	76	0.24	0.86	0.00	5.26
	Social Media	6	0.01	0.02	0.00	0.04
	Journal	84	0.15	0.33	0.00	1.72
	Letter	8	0.10	0.22	0.00	0.62
nonflu	**Prisoner**	36	0.05	0.07	0.00	0.21
	Mass Shooter	194	0.19	0.48	0.00	4.07
	Manifesto	20	0.12	0.32	0.00	1.32
	Blog	76	0.11	0.24	0.00	1.52
	Social Media	6	0.07	0.10	0.00	0.24
	Journal	84	0.31	0.65	0.00	4.07
	Letter	8	0.02	0.04	0.00	0.11
filler	**Prisoner**	36	0.03	0.07	0.00	0.30
	Mass Shooter	194	0.03	0.16	0.00	1.63
	Manifesto	20	0.05	0.15	0.00	0.66
	Blog	76	0.06	0.23	0.00	1.63
	Social Media	6	0.00	0.00	0.00	0.00
	Journal	84	0.02	0.08	0.00	0.58
	Letter	8	0.01	0.01	0.00	0.04
AllPunc	**Prisoner**	36	16.32	4.36	8.22	28.02

	Mass Shooter	194	25.52	12.69	3.12	94.74
	Manifesto	20	16.45	4.62	7.61	25.10
	Blog	76	30.58	15.62	8.33	94.74
	Social Media	6	15.86	9.93	3.12	32.21
	Journal	84	24.60	9.11	12.50	51.06
	Letter	8	17.12	5.21	10.56	26.81
Period	**Prisoner**	36	5.90	1.63	2.27	11.06
	Mass Shooter	194	8.84	6.58	0.00	36.84
	Manifesto	20	7.23	2.87	0.00	11.45
	Blog	76	9.30	8.60	0.00	36.84
	Social Media	6	5.78	3.84	0.00	10.76
	Journal	84	9.30	5.32	0.00	29.65
	Letter	8	6.06	1.95	3.36	10.26
Comma	**Prisoner**	36	4.73	1.96	0.00	9.74
	Mass Shooter	194	6.56	4.47	0.00	44.12
	Manifesto	20	4.49	2.71	0.36	10.00
	Blog	76	7.01	5.28	0.00	44.12
	Social Media	6	4.85	4.86	0.00	13.03
	Journal	84	6.91	4.01	0.00	20.59
	Letter	8	5.00	2.14	1.24	7.69
Colon	**Prisoner**	36	0.15	0.18	0.00	0.56
	Mass Shooter	194	2.56	4.98	0.00	22.22

	Manifest o	20	0.15	0.28	0.00	1.12
	Blog	76	6.00	6.57	0.00	22.22
	Social Media	6	0.34	0.55	0.00	1.40
	Journal	84	0.40	0.94	0.00	6.38
	Letter	8	0.28	0.38	0.00	1.08
SemiC	**Prisoner**	36	0.16	0.33	0.00	1.89
	Mass Shooter	194	0.05	0.15	0.00	1.06
	Manifesto	20	0.04	0.09	0.00	0.41
	Blog	76	0.03	0.10	0.00	0.65
	Social Media	6	0.07	0.11	0.00	0.23
	Journal	84	0.05	0.19	0.00	1.06
	Letter	8	0.15	0.17	0.00	0.36
QMark	**Prisoner**	36	0.41	0.59	0.00	2.83
	Mass Shooter	194	0.32	0.70	0.00	5.80
	Manifesto	20	0.32	0.45	0.00	1.63
	Blog	76	0.19	0.46	0.00	2.38
	Social Media	6	0.23	0.28	0.00	0.63
	Journal	84	0.40	0.72	0.00	3.36
	Letter	8	0.82	2.02	0.00	5.80
Exclam	**Prisoner**	36	0.13	0.30	0.00	1.29
	Mass Shooter	194	0.74	1.80	0.00	15.79
	Manifesto	20	0.22	0.40	0.00	1.31

	Blog	76	0.54	2.12	0.00	15.79
	Social Media	6	0.42	0.52	0.00	1.12
	Journal	84	1.14	1.75	0.00	8.20
	Letter	8	0.11	0.25	0.00	0.72
Dash	**Prisoner**	36	1.01	1.04	0.00	4.03
	Mass Shooter	194	1.50	2.90	0.00	25.32
	Manifesto	20	0.34	0.43	0.00	1.43
	Blog	76	3.21	4.03	0.00	25.32
	Social Media	6	0.45	0.50	0.00	0.98
	Journal	84	0.39	0.56	0.00	2.65
	Letter	8	0.57	0.76	0.00	2.26
Quote	**Prisoner**	36	0.94	0.90	0.00	2.95
	Mass Shooter	194	0.69	1.36	0.00	10.00
	Manifesto	20	0.45	0.49	0.00	1.21
	Blog	76	0.61	1.61	0.00	10.00
	Social Media	6	0.95	0.74	0.00	1.51
	Journal	84	0.81	1.35	0.00	6.67
	Letter	8	0.58	0.66	0.00	1.77
Apostro	**Prisoner**	36	1.58	1.30	0.00	5.47
	Mass Shooter	194	2.47	2.30	0.00	11.11
	Manifesto	20	2.10	2.66	0.00	10.00
	Blog	76	2.66	2.73	0.00	11.11

	Social Media	6	1.70	0.74	0.84	2.94
	Journal	84	2.44	1.89	0.00	8.82
	Letter	8	2.39	1.85	0.00	5.80
Parenth	**Prisoner**	36	0.57	0.88	0.00	5.09
	Mass Shooter	194	0.64	1.38	0.00	9.76
	Manifesto	20	0.49	0.62	0.00	1.77
	Blog	76	0.55	1.38	0.00	6.38
	Social Media	6	0.60	0.99	0.00	2.52
	Journal	84	0.77	1.57	0.00	9.76
	Letter	8	0.66	0.68	0.00	1.74
OtherP	**Prisoner**	36	0.74	1.02	0.00	4.99
	Mass Shooter	194	1.15	2.01	0.00	17.02
	Manifesto	20	0.61	0.74	0.00	2.65
	Blog	76	0.49	1.11	0.00	6.86
	Social Media	6	0.48	1.01	0.00	2.52
	Journal	84	2.00	2.60	0.00	17.02
	Letter	8	0.50	0.79	0.00	2.28
MFD Dic	**Prisoner**	36	2.11	1.00	0.00	4.92
	Mass Shooter	194	2.19	2.62	0.00	20.00
	Manifesto	20	3.01	4.19	0.00	20.00
	Blog	76	2.02	2.69	0.00	14.81
	Social Media	6	5.50	6.05	1.75	17.71

	Journal	84	1.87	1.41	0.00	6.45
	Letter	8	2.59	1.24	0.72	5.00
MFD General	**Prisoner**	36	0.24	0.27	0.00	1.03
	Mass Shooter	194	0.52	1.03	0.00	9.09
	Manifesto	20	0.43	0.54	0.00	1.77
	Blog	76	0.62	1.47	0.00	9.09
	Social Media	6	0.72	0.59	0.00	1.47
	Journal	84	0.46	0.62	0.00	2.27
	Letter	8	0.33	0.28	0.00	0.72
Harm	**Prisoner**	36	0.76	0.54	0.00	2.11
	Mass Shooter	194	0.86	2.22	0.00	20.00
	Manifesto	20	1.59	4.35	0.00	20.00
	Blog	76	0.76	1.88	0.00	14.81
	Social Media	6	3.36	6.53	0.33	16.67
	Journal	84	0.63	0.80	0.00	3.85
	Letter	8	0.63	0.60	0.00	1.89
Fairness	**Prisoner**	36	0.14	0.18	0.00	0.75
	Mass Shooter	194	0.10	0.36	0.00	2.94
	Manifesto	20	0.10	0.20	0.00	0.65
	Blog	76	0.11	0.43	0.00	2.94
	Social Media	6	0.16	0.24	0.00	0.55

	Journal	84	0.10	0.34	0.00	2.50
	Letter	8	0.13	0.17	0.00	0.47
Loyalty	**Prisoner**	36	0.33	0.30	0.00	1.47
	Mass Shooter	194	0.27	0.58	0.00	3.91
	Manifesto	20	0.39	0.49	0.00	1.86
	Blog	76	0.13	0.42	0.00	3.33
	Social Media	6	0.70	0.55	0.00	1.47
	Journal	84	0.28	0.58	0.00	3.23
	Letter	8	0.84	1.34	0.00	3.91
Authority	**Prisoner**	36	0.42	0.41	0.00	1.56
	Mass Shooter	194	0.22	0.42	0.00	2.78
	Manifesto	20	0.26	0.33	0.00	1.08
	Blog	76	0.14	0.28	0.00	1.39
	Social Media	6	0.30	0.39	0.00	1.04
	Journal	84	0.25	0.52	0.00	2.78
	Letter	8	0.49	0.53	0.00	1.63
Purity	**Prisoner**	36	0.12	0.16	0.00	0.57
	Mass Shooter	194	0.16	0.66	0.00	7.89
	Manifesto	20	0.18	0.35	0.00	1.35
	Blog	76	0.23	1.01	0.00	7.89
	Social Media	6	0.11	0.20	0.00	0.49
	Journal	84	0.10	0.23	0.00	1.23

Letter	8	0.04	0.05	0.00	0.11