Applying Natural Language Processing to Evaluate News Media Coverage of Bullying and Cyberbullying

Megan A. Moreno[1] · Aubrey D. Gower[1] · Heather Brittain[2] · Tracy Vaillancourt[2]

Abstract

Bullying events have frequently been the focus of coverage by news media, including news stories about teens whose death from suicide was attributed to cyberbullying. Previous work has shown that news media coverage is influential to readers in areas such as suicide, infectious disease outbreaks, and tobacco use. News media may be an untapped resource to promote bullying prevention messages, though current news media approaches to describing bullying and cyberbullying remain unexplored. The purpose of this study was to evaluate the current state of news media coverage of bullying and cyberbullying. A sample of newspaper articles covering bullying or cyberbullying across regional and national US newspapers from 6 recent years was identified. A content analysis using natural language processing was conducted with the Linguistic Inquiry and Word Count (LIWC) software program for key variables including affective, social, and cognitive processes. Evaluation included the percentage of words that represented Fear-based reporting such as alarmist words (e.g., epidemic, tragic), as well as words that represent Public Health-oriented messages such as prevention. A total of 463 newspaper articles met inclusion criteria, including 140 cyberbullying articles and 323 bullying articles. Findings indicated that cyberbullying articles scored higher on affective processes such as measures of anxiety (Mdn = 0.34) compared to bullying articles (Mdn = 0.22). A greater number of cyberbullying articles were Fear-based (41.4%) than were bullying articles (19.5%). An equivalent number of cyberbullying articles (50.0%) and bullying articles (49.8%) were Public Health-oriented. Findings may be used to collaborate with journalists toward optimizing prevention-oriented reporting.

Keywords Bullying · Cyberbullying · News media · Journalists · Natural language processing

Bullying is a preventable public health problem that can have long-lasting health, academic, and social consequences (McDougall and Vaillancourt 2015; Wolke and Lereya 2015). Bullying and cyberbullying have similarities and distinct elements. News media coverage of cyberbullying has included stories about individuals who have died from suicide, potentially influencing perceptions of parents or teens. Framing Theory (Entman 1993) describes how news media represent stories; cyberbullying stories may be represented with an over-emphasis on fear and tragedy. This study aims to assess text from bullying and cyberbullying newspaper articles to evaluate differences through the lens of Framing Theory.

Bullying and Cyberbullying: Common and Consequential

Current estimates support that traditional, or school-based, bullying affects 18–31% of youth and cyberbullying affects 7–15% of youth (CDC 2014). Bullying is defined using the Uniform Definition of Bullying (Gladden et al. 2014) to include factors such as unwanted aggression from peers and repetition. Within the Uniform Definition, cyberbullying is considered a context in which bullying may occur.

Cyberbullying has been associated with serious negative consequences such as depression, substance use, and school failure (Goebert et al. 2011; Schneider et al. 2012). Similar negative consequences are seen in cases of traditional bullying (Slonje and Smith 2008), but there is evidence that

✉ Megan A. Moreno
  mamoreno@pediatrics.wisc.edu; http://www.smahrtresearch.com

[1] Department of Pediatrics, University of Wisconsin-Madison, 2870 University Ave, Suite 200, Mailcode 9010, Madison, WI 53705, USA

[2] Faculty of Education, University of Ottawa, Ottawa, Canada

involvement in cyberbullying may contribute to depression and suicidality independent of traditional bullying (Bonanno and Hymel 2013; van Geel et al. 2014).

Cyberbullying: Similar to and Different from Bullying

Early studies of cyberbullying highlighted ways in which this type of bullying presented new challenges and distinct elements compared to traditional bullying. Studies have highlighted the 24/7 nature of cyberbullying, the potential for anonymity in bullying someone, and the capacity for a single bullying event to go viral (Patchin and Hinduja 2006; Slonje and Smith 2008).

However, other studies investigated both bullying and cyberbullying and found that these types of bullying often co-occur (del Rey et al. 2012; Landstedt and Persson 2014; Schneider et al. 2012) and that interventions targeting one type of bullying often have effects on the other (Palladino et al. 2016). A meta-analysis described significant relationships and theoretical overlap between cyberbullying and traditional bullying (Kowalski et al. 2014). The 2016 National Academies' report concluded that bullying and cyberbullying are more alike than dissimilar, and recommended that future studies should include both constructs (National Academies of Sciences 2016).

However, even as the scientific literature has moved toward viewing bullying and cyberbullying as more alike than dissimilar, perceptions of families may differ. A recent study focused on perceptions of parents and teens toward cyberbullying. The study findings indicated that their descriptions of a typical or example case of cyberbullying did not meet all criteria for the Uniform Definition of Bullying (Moreno et al. 2018). The most common components of the Uniform Definition described by participants included aggressive behavior, described by approximately 1/3 of participants. Participants also described cyberbullying involving anonymity, which lies outside the Uniform Definition. These findings suggest ways in which families may continue to see cyberbullying as distinct from traditional bullying.

One way in which families may derive impressions of bullying and cyberbullying experiences is from news media. Newspaper articles have suggested relationships between cyberbullying and mood disorders; an example article using this framing was a June 2015 Huffington Post article titled: Cyberbullying and Depression Go Hand in Hand (Pappas 2015). Individual cyberbullying stories have received national news coverage, some describing a death from suicide reported in the context of cyberbullying experiences (www.nobullying.com 2016). A previous study of cyberbullying news coverage found that news stories frequently focused on suicide, and concluded that news stories have the potential to normalize suicide as a response to

cyberbullying (Young et al. 2017). Youth and families may be exposed to news coverage focused on tragic cases of cyberbullying, which may affect public perceptions or influence social norms around cyberbullying.

News Media and Framing Theory

This study's theoretical approach is centered on Framing Theory. Framing Theory refers to the process of selecting certain aspects of reality and makes them more salient, while leaving other aspects absent (Entman 1993). It is through this selection process that journalists promote a specific framing of an issue, leading the audience to a particular interpretation. The rationale for framing by news media is to reduce the complexity of issues by presenting them in easy to understand interpretive packages (Kim and Willis 2007). Through framing, news media provide a perspective on the nature of the leading problems in their community, and illustrate what people should fear and what they should care about (Clegg Smith 2014).

Although little is known about the impact on families or communities following news media coverage of bullying and cyberbullying, this coverage has been criticized as alarmist and oversimplified (Tenore 2013). News coverage may use framing to support that bullying or cyberbullying is something to fear, or that cannot be prevented. In contrast, news media may use framing to describe bullying or cyberbullying as something that can be prevented and addressed.

Previous work supports the influence of news media on the lay public's perceptions and views. In one study, investigators showed participants different news media representations about mental health and found that views about recovery, stigma, and attitudes toward mental health were influenced by news media language (Corrigan et al. 2013). A previous study concluded that many people rely on news media at least as much as they rely on their healthcare providers for information about illness and disease (Donelle et al. 2005). These studies illustrate the role of news media in influencing readers' perceptions and views toward critical health topics, often via framing approaches.

While little is known about the impact of bullying news media coverage on families, a related situation may be to consider the impact of news stories in cases of suicide. Previous studies have shown that people bereaved by suicide are often distressed by news media reporting that dramatizes the event or chastises the victim (Chapple et al. 2013; Cover 2012). A further risk in news coverage of suicide is to influence similar behavior, known as copycat suicides (Gould et al. 2003) or the Werther effect (Phillips 1974). Studies have shown increases in suicides following a news report of suicide (Bollen and Phillips 1982; Phillips 1974). The evidence from these studies supports the influence of news media and the need for understanding how bullying is portrayed via these outlets.

Text-Based Analysis

Studies of text have allowed researchers to understand dimensions of communication and appeal in written messages, including elements that make up the framing of these messages. A common approach in text-based analysis is to understand if messages are designed to appeal to emotion versus appealing to rationality. Messages relating to emotion often include emotional words (e.g., happy, angry), and those relating to rationality may include facts or statistics. These different framings may have different appeal or influence on readers. One example study evaluated potential intervention messages around smoking cessation to achieve a balance between emotional and rationale appeal (Khalil et al. 2018).

Beyond the manual analysis of text, software programs have been developed to automatically analyze the content. These programs evaluate message content based on categories that were previously developed through extensive traditional manual coding. A validated program that is commonly used is Linguistic Inquiry and Word Count (LIWC). LIWC is built upon the premise that the words that people use can provide rich information about their beliefs, fears, thinking patterns, relationships and views. The LIWC program can objectively and quantitatively identify these informational features of that text's message.

An Important Gap in the Literature

In order to understand the potential impact of news media about bullying, the current state of the content and thus framing of these news media articles must be understood. Elements of cyberbullying that are perceived as unique may lead to news articles using a different tone, word choice or emotionality compared to traditional bullying. If present, these differences in newspaper articles represent different framing, which could influence how families interpret or react to cyberbullying compared to bullying.

This study seeks to understand whether specific types of framing are present in bullying and cyberbullying newspaper articles. Given that previous cyberbullying newspaper articles have focused on tragic cases or death from suicide, this study will develop and test a Fear-based dictionary within LIWC to evaluate that framing approach. This study defined this Fear-based framework as focusing on negative consequences or presenting bullying situations as hopeless or inevitable. Given researchers' and policy makers' interest in prevention and public health approaches for bullying, this study will develop and test a Public Health-oriented dictionary within LIWC to evaluate that framing approach. This study defined Public Health-oriented as having an emphasis on prevention, community engagement or

practical approaches. Thus, the specific aims and hypotheses of this study are summarized below:

1. Specific Aim 1: To compare bullying and cyberbullying newspaper articles using text-based analysis using a validated tool and a priori related dictionaries

   Hypothesis: Cyberbullying newspaper articles will be more likely to use emotional language or language around harm compared to bullying articles.

2. Specific Aim 2: To develop and test a Fear-based and Public Health-oriented approach and LIWC dictionaries to apply to bullying and cyberbullying newspaper articles

   Hypothesis: Cyberbullying newspaper articles will be more likely to use a Fear-based reporting approach compared to bullying articles.

Summary and Purpose

Previous literature supports the role of news media as influential on perceptions and views about health-related topics. These studies support the need to understand the framing and communication style in which bullying and cyberbullying portrayed via the news media. As per the National Academies of Sciences, Engineering, and Medicine (2016) report, bullying and cyberbullying should be studied in tandem rather than separately. However, previous studies also suggest that cyberbullying is a newer phenomenon that may have unique elements. It remains unclear whether the current state of news media coverage of bullying uses particular framing elements grounded in emotional compared to rationale language. Further, it is unclear whether newspaper articles utilize framing grounded in Fear-based or Public Health-oriented word choices and further whether this differs by whether the article topic is bullying or cyberbullying. The purpose of this study was twofold: to conduct a content analysis of text-based news media coverage of bullying and cyberbullying, and to develop and apply a Fear-based and Public Health-oriented framework to evaluate this news media coverage.

Method

This content analysis study used natural language processing on a sample of published newspaper articles related to bullying and cyberbullying from 2012 to 2017. As content analysis of news media coverage related to bullying has not previously been conducting using this methodologic approach, the nature of this study is exploratory.

Newspaper Selection

The goal of newspaper selection was to include a geographically diverse selection of both regional and national newspapers. Inclusion criteria limited newspapers to those in the USA, those having availability on ProQuest, and those having high readership numbers, defined as a daily readership of 150,000 or higher. This resulted in 13 newspapers eligible for investigation. Stratified sampling was used to select newspapers to vary geographically to include approximately half regional and half national sources. Four national newspapers that met readership criteria were included: The Washington Post, New York Times, Los Angeles Times and USA Today (center 1998). For regional newspapers, if two newspapers had high readership but were in a similar geographic area (such as Philadelphia and Pittsburgh), random selection was used to pick one for inclusion in the study. Following this selection process, the final selected newspapers included the following: The Washington Post, New York Times, Chicago Tribune, Los Angeles Times, USA Today, Arizona Republic, Tampa Bay Times, Pittsburgh Post-Gazette, and San Jose Mercury (The Mercury).

Newspaper Article Selection

The goal of the newspaper article search strategy was to identify a sample of published articles on bullying and cyberbullying from the past 6 years. The online electronic database ProQuest: US News-stream was used to search for newspaper articles. ProQuest's search engine sorts articles by relevance using an algorithm that is based on the number of times the search terms appear in an article. Using the advanced search option, two separate searches were conducted to identify potentially eligible articles; for bullying articles the search term bullying was used, while the search for cyberbully articles used the search term cyberbullying. Search criteria limited articles to full-text newspaper articles published in selected newspapers with publication date between January 1, 2012 and October 31, 2017.

Based on the search strategy, potentially eligible newspaper articles were downloaded from both initial searches ($n =$ 1477), and then each article was read and evaluated to apply inclusion criteria. Two investigators conducted newspaper article review, with one serving as primary reviewer and the other to review subsets of articles for reliability. Questionable articles during the review process were flagged and discussed to reach consensus. Inclusion criteria limited articles to those for which the article topic was exclusively focused on bullying or cyberbullying (did not include articles that focused on both simultaneously) given our interest in comparing bullying articles and cyberbullying articles. Inclusion criteria also limited newspaper articles to those focused on adolescents and young adults aged 10–25 years; this inclusion criteria meant that news articles were included that addressed bullying among youth in elementary, middle and high school as well as bullying events in college. Articles that reported on the same bullying event were included, though identical articles in different newspapers (such as a Reuter's article published in several papers) were excluded.

Linguistic Inquiry and Word Count

Natural language processing was conducted applying the Linguistic Inquiry and Word Count (LIWC) program (Linguistic Inquiry and Word Count: LIWC2015, Pennebaker Conglomerates Inc., Austin, TX). LIWC is a computerized text-analysis software that counts the frequency of words and word stems to understand elements such as structure, process, emotional, and cognitive components in written text (Pennebaker et al. 2015). This software program builds on several decades of research to understand narrative voice in health (Berry and Pennebaker 1993; Francis and Pennebaker 1992; Pennebaker 1993; Pennebaker and King 1999; Pennebaker et al. 1997; Rosenberg et al. 2002).

LIWC Dictionaries The LWIC program calculates standardized composites for summary scores of Analytic, Clout, Authentic, and Tone, which are then converted to percentiles, based on large samples (LIWC 2016). During LIWC analysis of a document, every word is then compared with dictionaries of up to 74 dimensions. LIWC calculates a percentage of total words falling into different dictionaries (LIWC 2016). Examples of categories of evaluation include summary language variables, such as words per sentence or words > 6 letters, linguistic dimensions such as pronouns, adverbs or conjunctions, and psychological processes.

Psychological processes are a broad category including multiple dictionaries. Dictionaries included Affective Processes such as positive or negative emotion, as well as Social, Cognitive, Perceptual, and Biological Processes. Additional dictionaries include Drives, Time Orientations, Relatively, Personal Concerns, and Informal Language. These dictionaries (Table 2) were developed by a rigorous process in which groups of judges reviewed 2000 words or word stems and determined how the reviewed words related to specific categories. The LIWC program also allows for construction of new dictionaries.

Existing LIWC dictionaries were developed through a rigorous process including word collection, judge rating, base rate analyses, candidate word list generation, refinement and psychometric evaluation (Pennebaker et al. 2015). LIWC has been validated for content and construct validity (Francis and Pennebaker 1992; Stirman and Pennebaker 2001). Interrater reliability discrimination of categories has been found to range from 86 to 100% depending on the dimension being assessed, supporting content validity.

Measures

LIWC Dictionaries Given the sheer number of dictionaries within the psychological processes category, LIWC dictionaries were selected a priori that had salience from previous studies of bullying and cyberbullying, or that aligned with the specific aims of our study. First, based on associations between bullying experiences and affective and emotional experiences (Ttofi et al. 2011), the Affective Processes dictionary was selected. Further, affective processes have previously been evaluated using LIWC (Khalil et al. 2018; Stirman and Pennebaker 2001; Wardecker et al. 2017) and identified as influential dimensions of language. These variables were also specified in the first aim of this study related to emotional language and our hypotheses. Affective Processes included words related to positive and negative emotion, anxiety, anger, and sadness.

Second, given that bullying and cyberbullying occur in a social context involving peers (Gladden et al. 2014), the Social Processes dictionary was selected. Social Processes are frequently assessed in studies using LIWC to understand the presence of processes such as support (Bantum and Owen 2009; Heckler et al. 2014). Social Processes included words such as social, family, friend, female, and male.

Third, because it was hypothesized that journalists may describe bullying and cyberbullying with different levels of certainty or insight given that cyberbullying is a newer and thus less well understood phenomenon, the Cognitive Processes dictionary was included. Examples of Cognitive Processes words included insight, causation, discrepancy, tentative, certain, and differing. Fourth, given the link between bullying experiences and health (John et al. 2018; Karanikola et al. 2018), the Biological Processes dictionary was included, which included words describing health as well as references to death. Fifth, as bullying is defined as involving power imbalances (Gladden et al. 2014; Olweus 2013), the Drives dictionary was selected, which included words related to affiliation, achieve, power, reward, and risk. Finally, it was posited that bullying and cyberbullying news articles may have a different focus regarding typically prevention reporting in a future-oriented manner, versus potentially reporting a tragedy in a past-oriented manner, the dictionary describing Time Orientations was included.

Fear-Based and Public Health-Oriented The second study goal was to develop and apply a LIWC approach to identify articles that were primarily Fear-based or Public Health-oriented, as well as articles incorporating both frameworks. A modified Delphi method (Dalkey 1969) was applied to develop, test, and refine these two categories applying processes derived from previous LIWC studies that created new dictionaries (Khalil et al. 2018). First, investigators identified initial criteria and keywords to represent Fear-based and Public Health-oriented reporting based on a review of the literature and reading a random sample of newspaper articles that met inclusion criteria but were from newspapers not included in the study. Second, while reviewing news articles during the eligibility assessment, investigators identified example articles that represented focused examples of key constructs involved in Fear-based and Public Health-oriented reporting ($n = 10$). These articles were then reviewed and 4 were selected through a process of discussion and investigator consensus across 3 investigators.

Investigators then identified a group of 10 stakeholders including pediatricians, researchers, parents, and adolescents as our Delphi review panel. Stakeholders were asked to review and categorize each of the four example newspaper articles by applying and revising the initial proposed keywords representing Fear-based and Public Health-oriented. Stakeholders were then asked to identify additional words which they believed aided in their decision to categorize an example article as Fear-based and Public Health-oriented.

For the second round of the Delphi procedure, feedback was incorporated and words proposed by stakeholders were added that would represent Fear-based and Public Health-oriented. Stakeholders were asked for a second round of review and feedback. From the second round of feedback, words with consensus across stakeholders and researchers were included to create the final dictionary.

Based on the Delphi process, a dictionary was created to evaluate article content as either Fear-based or Public Health-oriented. For Fear-based, example words included victim, threatened and harassed. For Public Health-oriented, example words included prevention, community, intervention and outreach. To determine categorization of the articles that contained both Fear-based and Public-Health reporting, a median split of differences between reporting type from all articles was used as a threshold (Dzindolet and Pierce 2005; Sumner et al. 2012). The median threshold was established as 0.155. Newspaper articles were categorized by which type of reporting was more prevalent, so long as the absolute difference was greater than the threshold. Article in which the difference in percentage of Fear-based and Public Health-oriented reporting was higher than the threshold (0.155), the article was categorized as either Fear-based or Public Health-oriented. Articles that contained equal percentages of Fear-based and Public Health-oriented reporting, or those in which the difference was lower than the threshold, were categorized as mixed. Articles that did not register any percentage for either of these categories were excluded from final analysis.

Analysis

All newspaper article text was imported and evaluated using the LIWC program (Linguistic Inquiry and Word Count: LIWC2015, Pennebaker Conglomerates Inc., Austin, TX).

Output scores from LIWC were assessed for distribution to determine appropriate analyses for comparison. Given that many variables had a positively skewed distribution and high kurtosis, indicating non-normality of the distribution (i.e., many articles had low proportions of these words and a few had a greater proportion), non-parametric tests were used. The Mann-Whitney U test, similar to the independent t test for non-parametric data, was used to make univariate comparisons between bullying and cyberbullying newspaper articles across variables of interest. The test is used to compare distributional differences of spread (sensitive to differences in central tendency) between independent groups (Howell 2012). Plots of the distributions were examined to ensure similar shape so that differences could be attributed to median differences between article types. Any statistically significant differences indicate that the medians of the variable differ by article type, that is bullying and cyberbullying articles, were different.

The effect size of η^2 was reported, which is interpreted as the proportion of variance in the LIWC outcome attributable to differences between bullying and cyberbullying articles. Due to the number of comparisons, the Benjamini-Hochberg (BH) correction for multiple testing (Benjamini and Hochberg 1995) was used. Due to the categorical classifications of Fear-based and Public Health-oriented reporting, a Chi-square test was performed comparing bullying and cyberbullying articles on these two dimensions.

Results

Summary of Newspaper Articles

A total of 463 newspaper articles met inclusion criteria: 140 cyberbullying articles and 323 bullying articles. Among cyberbullying articles, 477 articles were reviewed for eligibility, 309 (64.8%) were excluded based on topic other than youth cyberbullying, and 28 (5.8%) were duplicate articles. Of the cyberbullying articles, 38.6% were from a regional newspaper and 61.4% were from a national newspaper. Among bullying articles, 1000 newspaper articles were reviewed for eligibility, 624 (62.4%) were excluded for a focus other than youth bullying and 53 (5.3%) were duplicates. The bullying articles had a similar distribution for regional and national. Table 1 illustrates descriptive information about news articles.

LIWC Dictionary: Bullying Compared to Cyberbullying

The percentages of words by category in bullying and cyberbullying newspaper articles are shown in Table 2. Correcting for multiple testing results indicated that none of the summary language variables were statistically

Table 1 Descriptive information about news media articles and sources

	Cyberbullying	Bullying
Number of articles	140	323
By year		
2012	28 (20.0%)	71 (22.0%)
2013	46 (32.9%)	77 (23.8%)
2014	26 (18.6%)	72 (22.3%)
2015	22 (15.7%)	45 (13.9%)
2016	14 (10.0%)	36 (11.1%)
2017	4 (2.9%)	22 (6.8%)
By source		
Arizona Republic	10 (7.1%)	42 (13.0%)
Chicago Tribune	29 (20.7%)	72 (22.3%)
New York Times	12 (8.6%)	36 (11.1%)
LA Times	9 (6.4%)	23 (7.1%)
Washington Post	32 (22.9%)	58 (18.0%)
San Jose Mercury	18 (12.9%)	16 (5.0%)
USA Today	4 (2.9%)	14 (4.3%)
Tampa Bay Times	20 (14.2%)	34 (10.5%)
Pittsburgh Post-Gazette	6 (4.3%)	28 (8.7%)
Type of paper		
Regional	54 (38.6%)	120 (37.2%)
National	86 (61.4%)	203 (62.8%)

significantly different between article topic. In evaluation of affective processes, LIWC scores of anxiety-related words were higher for cyberbullying articles (Mdn = 0.34) than for bullying articles (Mdn = 0.22), $U = 27{,}059.00$, $Z = 3.39$, $p = .001$, $\eta^2 = .025$. The same was found for scores for words related to anger ($Mdn_{cyberbullying} = 0.86$; $Mdn_{bullying} = 0.61$), $U = 27{,}418.00$, $Z = 3.64$, $p < .001$ $\eta^2 = .029$.

For the Social Processes dictionary, none of the variables were significantly different by article topic. In evaluating Cognitive Processes, differences were found in both cause and tentative word use. Specifically, cyberbullying articles had higher percentages of cause-related words than bullying articles ($Mdn_{cyberbullying} = 1.77$; $Mdn_{bullying} = 1.59$), $U = 26{,}150.00$, $Z = 2.68$, $p = .007$, $\eta^2 = .016$. The same was found for the percentages of tentative words ($Mdn_{cyberbullying} = 2.11$; $Mdn_{bullying} = 1.75$), $U = 26{,}405.00$, $Z = 2.87$, $p = .004$, $\eta^2 = .018$.

Following the correction for multiplicity, the Biological Process variable of health was not statistically different across article topic. However, although bullying and cyberbullying articles included few mentions of death ($Mdn_{cyberbullying} = 0.20$; $Mdn_{bullying} = 0.00$), the difference in the distributions was statistically significant, $U = 28{,}035.00$, $Z = 4.45$, $p < .001$, $\eta^2 = .043$.

Within the Drives dictionary, two of the six categories were statistically significant. More affiliation words were found in

Table 2 Results from Linguistic Inquiry and Word Count (LIWC) analysis of bullying and cyberbullying news articles

LIWC dictionary	LIWC terms	Articles median percentile scores		$p_{adjusted}$
Summary variables		Bullying	Cyberbullying	
Analytic	–	89.24	87.80	.211
Clout	–	81.81	83.44	.773
Authentic	–	16.57	16.22	.773
		Articles median proportion scores		$p_{adjusted}$
Affective processes		Bullying	Cyberbullying	
Positive emotions	Love, nice, and sweet	1.97	1.99	.773
Negative emotions	Hurt, ugly, and nasty	2.04	2.28	.165
Anxiety	Worried and fearful	0.22	0.34	*.008*
Anger	Hate, kill, and annoyed	0.61	0.86	*< .001*
Sad	Crying, grief, and sad	0.26	0.20	.115
Social processes				
Social	Mate, talk, and they	12.40	13.13	.773
Family	Daughter, dad, and aunt	0.54	0.52	.434
Friend	Buddy and neighbor	0.25	0.27	.603
Female	Girl, her, and mom	0.74	1.16	.112
Male	Boy, his, and dad	0.93	0.72	.080
Cognitive processes				
Insight	Think and know	1.91	1.92	.603
Causal	Because and effect	1.59	1.77	*.038*
Tentative	Maybe and perhaps	1.75	2.11	*.023*
Certainty	Always and never	0.85	0.83	.840
Differentiation	Has not, but, and else	2.57	2.65	.528
Biological processes				
Health	Clinic, flu, and pill	0.40	0.30	.083
Personal concerns				
Death	Bury, coffin, and kill	0.00	0.20	*< .001*
Drives				
Affiliation	Ally, friend, and social	2.33	2.81	*.008*
Power	Superior and bully	4.93	4.42	.112
Reward	Take, prize, and benefit	0.74	0.87	.603
Risk	Danger and doubt	0.70	0.83	*.038*
Time orientations				
Past focus	Ago, did, and talked	5.35	4.30	.083
Present focus	Today, is, and now	6.63	7.53	.083
Future focus	May, will, and soon	0.66	0.64	.727

Adjusted *p* value from Mann-Whitney *U* test, significant values in italics

cyberbullying newspaper articles (Mdn = 2.81) than bullying articles (Mdn = 2.33), $U = 26{,}885.00$, $Z = 3.23$, $p = .001$, $\eta^2 = .023$. Words involving risk were also found to be more prominent in the cyberbullying articles (Mdn = 0.83) than the bullying articles (Mdn = 0.70), $U = 26{,}117.00$, $Z = 2.65$, $p = .008$, $\eta^2 = .015$. Articles did not vary by time orientations.

Fear-Based Versus Public Health-Oriented

Bullying and cyberbullying articles demonstrated an association between article topic and article focus. An approximately

equal number of bullying (49.8%) and cyberbullying articles (50.0%) were categorized as Public Health-oriented. However, a greater number of cyberbullying articles were Fear-based (41.4%) than were bullying articles (19.5%), $\chi^2(2, N = 460) = 23.196$. Table 3 illustrates these findings.

Discussion

This study used a validated natural language processing approach which compared percentages of word counts within

Table 3 Comparison of Fear-based and Public Health-oriented news media articles on bullying and cyberbullying

	Fear-based	Public Health-oriented	Mixed or equal Fear-based and Public Health-oriented
Bullying, $n = 323$	19.5% ($n = 63$)	49.8% ($n = 161$)	30.6% ($n = 99$)
Cyberbullying, $n = 137$	42.3% ($n = 58$)	51% ($n = 70$)	4.4% (n = 9)

3 articles were excluded as they did not register as fear/public health/mix based on analysis

dictionaries to evaluate the current state of news media for bullying and cyberbullying. The theoretical approach of Framing Theory was used in this study, and was applied to LIWC dictionaries, as well as a novel Fear-based and Public Health-oriented framework, to evaluate newspaper articles. Study hypotheses that cyberbullying articles would have more frequent use of language around harm, as well as more frequent Fear-based framework reporting, were found to be supported. The study results suggest that newspaper coverage of cyberbullying does present unique framing in comparison to bullying reporting. This study presents unique data toward understanding the type of content available via news media to parents, adolescents, and families.

A first finding was areas in which reporting approach for bullying and cyberbullying were alike. Both types of articles incorporated emotional language, including both positive and negative emotion words. Both article types used social orientation of placing bullying in the context of family and peer processes. This finding is consistent with bullying taking place in a person by situation by context approach (Ortega-Ruiz and Nunez 2012). Both types of newspaper articles included similar frequency of terms related to power and reward, consistent with key constructs in the Uniform Definition of Bullying (Gladden et al. 2014).

A second finding was ways in which news media bullying reporting differed from cyberbullying. Cyberbullying articles were more likely to include words referring to anxiety and anger. Further, cyberbullying articles were more likely to use language related to death. These language choices may be related to cyberbullying articles describing tragic bullying cases. The website www.nobullying.com includes an article describing "The Top Six Unforgettable Cyberbullying Cases Ever," describing teen deaths from suicide. What remains unclear is whether this finding is linked to journalists focusing on the facts of the stories they are reporting, or whether journalists are using Framing Theory to draw attention to cyberbullying as a particularly fear-based topic. Future studies should consider journalists' views and experiences in reporting on cyberbullying, and whether editorial or societal pressures influence reporting style.

Finally, applying a new Fear-based compared to Public Health-oriented framework identified that cyberbullying newspaper articles were more likely to have a Fear-based approach compared to bullying articles. This finding supports the proposed

hypothesis that cyberbullying articles were more likely to use Fear-based words. This finding is also consistent with cyberbullying articles often focusing on tragic cases or even suicide. However, bullying articles and cyberbullying articles were approximately equal on Public Health-oriented content. These findings together may suggest that a single newspaper article on cyberbullying may incorporate both approaches, such as describing a case of suicide alongside tips for prevention.

Limitations

Limitations to this study include the use of specific search terms to identify newspaper articles which may not have captured all content that describes bullying and cyberbullying. It is possible that there are different dimensions to language differences across newspaper articles that describe teasing, harassment, or other related terms. A second important limitation is that LIWC content analysis provides a systematic quantification of news article text which does not factor in overall context. While a limitation to the LIWC tool, it must be balanced against the strength of LIWC to provide an objective and quantitative assessment of large quantities of text. The LIWC assessment assumes that the words within the dictionaries are appropriate proxies for article meaning, though there may be other interpretations of certain words in certain contexts that LIWC could not distinguish. However, the dictionaries used by LIWC have been well-supported by previous literature. Third, LIWC uses an automated programmed process; it captures frequency of words in a category which has limitations compared to manual coding. Thus, appreciation of words used in jest or sarcasm is interpreted at face value. However, this was unlikely to be a significant issue in this study of news media. Fourth, this study focused on text-based news media, which means that television and some forms of online content were excluded. Finally, this study was limited to a selection of newspaper articles in English in the USA over a specific time period, and findings may not apply to other newspapers, countries, or time periods.

Implications and Conclusion

Implications of this study include the possibility to collaborate with journalists to improve news media coverage of bullying

and cyberbullying toward the goal of enhancing prevention. This potential implication may be well illustrated in a different yet related context of suicide reporting in which public health practitioners and researchers collaborated with journalists to develop reporting standards with an emphasis on prevention. A 2017 study examined newspaper articles pre- and post-implementation of media guidelines around responsible suicide reporting. Study findings indicated that news reporting in the follow-up period was less sensationalistic and more prevention focused (Roskar et al. 2017). This study illustrates an effective example of implementing media reporting guidelines, as well as content analysis of news media pre- and post-intervention.

Findings from this study may inform efforts to improve news media reporting on bullying, particularly cyberbullying. For example, if bullying news media stories highlighted successful school-based interventions for bullying, this may raise awareness and provide parents with tools to advocate for such programs in their schools. Leveraging findings such as these to collaborate with journalists may provide new opportunities to enhance prevention.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflicts of interest.

Ethical Approval This article does not contain any studies with human participants performed by any of the authors.

Informed Consent No informed consent was conducted as there were no human participants.

References

Bantum, E. O., & Owen, J. E. (2009). Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychological Assessment, 21*, 79–88. https://doi.org/10.1037/a0014643.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological, 57*, 289–300.

Berry, D. S., & Pennebaker, J. W. (1993). Nonverbal and verbal emotional expression and health. *Psychotherapy and Psychosomatics, 59*, 11–19.

Bollen, K. A., & Phillips, D. P. (1982). Imitative suicides: a national study of the effects of television news stories. *American Sociological Review, 47*, 802–809.

Bonanno, R. A., & Hymel, S. (2013). Cyber bullying and internalizing difficulties: above and beyond the impact of traditional forms of bullying. *Journal of Youth and Adolescence, 42*, 685–697. https://doi.org/10.1007/s10964-013-9937-1.

CDC. (2014). Youth risk behavior surveillance - United States. Retrieved from https://www.cdc.gov/mmwr/pdf/ss/ss6304.pdf. Accessed 3 Jan 2019.

center, P. r. (1998). Local versus National Papers. Retrieved from https://www.journalism.org/1998/07/13/local-versus-national-papers/. Accessed 3 Jan 2019.

Chapple, A., Ziebland, S., Simkin, S., & Hawton, K. (2013). How people bereaved by suicide perceive newspaper reporting: qualitative study. *The British Journal of Psychiatry, 203*, 228–232. https://doi.org/10.1192/bjp.bp.112.114116.

Clegg Smith, K. (2014). News media analysis: influence on community health behaviors and practices. In S. M. Albert (Ed.), *Methods for community public health research: Integrated and engaged approaches*. Springer.

Corrigan, P. W., Powell, K. J., & Michaels, P. J. (2013). The effects of news stories on the stigma of mental illness. *The Journal of Nervous and Mental Disease, 201*, 179–182. https://doi.org/10.1097/NMD.0b013e3182848c24.

Cover, R. (2012). Mediating suicide: print journalism and the categorization of queer youth suicide discourses. *Archives of Sexual Behavior, 41*, 1173–1183. https://doi.org/10.1007/s10508-012-9901-2.

Dalkey, N. (1969). The Delphi method: an experimental study of group opinion. Retrieved from https://www.rand.org/content/dam/rand/pubs/research_memoranda/2005/RM5888.pdf. Accessed 3 Jan 2019.

del Rey, R., Elipe, P., & Ortega-Ruiz, R. (2012). Bullying and cyberbullying: overlapping and predictive value of the co-occurrence. *Psicothema, 24*, 608–613.

Donelle, L., Hoffman-Goetz, L., & Clarke, J. N. (2005). Ethnicity, genetics, and breast cancer: media portrayal of disease identities. *Ethnicity & Health, 10*, 185–197. https://doi.org/10.1080/13557850500120751.

Dzindolet, M., & Pierce, L. (2005). Using a linguistic analysis tool to detect deception. In *Proceedings of the human factors and ergonomics society 49th annual meeting*. https://doi.org/10.1177/154193120504900374.

Entman, R. M. (1993). Framing: toward clarification of a fractured paradigm. *Journal of Communication, 43*, 51–58.

Francis, M. E., & Pennebaker, J. W. (1992). Putting stress into words: the impact of writing on physiological, absentee, and self-reported emotional well-being measures. *American Journal of Health Promotion, 6*, 280–287.

Gladden, R. M., Vivolo-Kantor, A. M., Hamburger, M. E., & Lumpkin, C. D. (2014). Bullying surveillance among youths: uniform definitions for public health and recommended data elements. Retrieved from Atlanta, GA: http://www.cdc.gov/violenceprevention/pdf/bullying-definitions-final-a.pdf. Accessed 3 Jan 2019.

Goebert, D., Else, I., Matsu, C., Chung-Do, J., & Chang, J. Y. (2011). The impact of cyberbullying on substance use and mental health in a multiethnic sample. *Maternal and Child Health Journal, 15*, 1282–1286. https://doi.org/10.1007/s10995-010-0672-x.

Gould, M. S., Greenberg, T., Velting, D. M., & Shaffer, D. (2003). Youth suicide risk and preventive interventions: a review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 386–405. https://doi.org/10.1097/01.Chi.0000046821.95464.Cf.

Heckler, E. B., Dubey, G., McDonald, D. W., Poole, E. S., Li, V., & Eikey, E. (2014). Exploring the relationship between changes in weight and utterances in an online weight loss forum: a content and correlational analysis study. *Journal of Medical Internet Research, 16*, e254.

Howell, D.C. (2012). Statistical methods for psychology (PSY 613 Qualitative Research and Analysis in Psychology), 8th Edition. Boston: Cengage Learning.

John, A., Glendenning, A. C., Marchant, A., Montgomery, P., Stewart, A., Wood, S., . . ., Hawton, K. (2018). Self-harm, suicidal behaviours, and cyberbullying in children and young people: systematic review. *Journal of Medical Internet Research, 20*, e129. https://doi.org/10.2196/jmir.9044.

Karanikola, M. N. K., Lyberg, A., Holm, A. L., & Severinsson, E. (2018). The association between deliberate self-harm and school bullying victimization and the mediating effect of depressive symptoms and self-stigma: a systematic review. *BioMed Research International, 2018*, 4745791. https://doi.org/10.1155/2018/4745791.

Khalil, G. E., Calabro, K. S., Crook, B., Machado, T. C., Perry, C. L., & Prokhorov, A. V. (2018). Validation of mobile phone text messages for nicotine and tobacco risk communication among college students: a content analysis. *Tobacco Prevention and Cessation, 4*. https://doi.org/10.18332/tpc/84866.

Kim, S. H., & Willis, L. A. (2007). Talking about obesity: news framing of who is responsible for causing and fixing the problem. *Journal of Health Communication, 12*, 359–376. https://doi.org/10.1080/10810730701326051.

Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin, 140*, 1073–1137. https://doi.org/10.1037/a0035618.

Landstedt, E., & Persson, S. (2014). Bullying, cyberbullying, and mental health in young people. *Scandinavian Journal of Public Health, 42*, 393–399. https://doi.org/10.1177/1403494814525004.

LIWC. (2016). Dictionaries on LIWC. Retrieved from http://liwc.wpengine.com/compare-dictionaries/. Accessed 3 Jan 2019.

McDougall, P., & Vaillancourt, T. (2015). Long-term adult outcomes of peer victimization in childhood and adolescence: pathways to adjustment and maladjustment. *American Psychologist, 70*, 300.

Moreno, M. A., Suthamjariya, N., & Selkie, E. (2018). Stakeholder perceptions of cyberbullying cases: application of the uniform definition of bullying. *The Journal of Adolescent Health, 62*, 444–449. https://doi.org/10.1016/j.jadohealth.2017.11.289.

National Academies of Sciences, E. a. M. (2016). Preventing bullying science, policy, and practice. Retrieved from Washington, DC.

Olweus, D. (2013). School bullying: development and some important challenges. *Annual Review of Psychology, 9*, 751–780.

Ortega-Ruiz, R., & Nunez, J. C. (2012). Bullying and cyberbullying: research and intervention at school and social contexts. *Psicothema, 24*, 603–607.

Palladino, B. E., Nocentini, A., & Menesini, E. (2016). Evidence-based intervention against bullying and cyberbullying: evaluation of the NoTrap! Program in two independent trials. *Aggressive Behavior, 42*, 194–206. https://doi.org/10.1002/ab.21636.

Pappas, S. (2015). Cyberbullying and depression go hand in hand according to new study. Huffington Post. Retrieved from https://www.huffpost.com/entry/cyberbullying-depression_n_7644934. Accessed 3 Jan 2019.

Patchin, J. W., & Hinduja, S. (2006). Bullies move beyond the schoolyard a preliminary look at cyberbullying. *Youth Violence and Juvenile Justice, 4*, 148–169.

Pennebaker, J. W. (1993). Putting stress into words: health, linguistic, and therapeutic implications. *Behaviour Research and Therapy, 31*, 539–548.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: language use as an individual difference. *Journal of Personality and Social Psychology, 77*, 1296–1312.

Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*, 863–871.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Balckburn, K. (2015). The development and psychometric properties of LIWC2015. Retrieved from Austin, TX.

Phillips, D. P. (1974). The influence of suggestion on suicide: substantive and theoretical implications of the Werther effect. *American Sociological Review, 39*, 340–354.

Rosenberg, H. J., Rosenberg, S. D., Ernstoff, M. S., Wolford, G. L., Amdur, R. J., Elshamy, M. R., . . ., Pennebaker, J. W. (2002). Expressive disclosure and health outcomes in a prostate cancer population. *International Journal of Psychiatry in Medicine, 32*, 37–53.

Roskar, S., Tancic Grum, A., Postuvan, V., Podlesek, A., & De Leo, D. (2017). The adaptation and implementation of guidelines for responsible media reporting on suicide in Slovenia. *Zdr Varst, 56*, 31–38. https://doi.org/10.1515/sjph-2017-0005.

Schneider, S. K., O'Donnell, L., Stueve, A., & Coulter, R. W. (2012). Cyberbullying, school bullying, and psychological distress: a regional census of high school students. *American Journal of Public Health, 102*, 171–177. https://doi.org/10.2105/AJPH.2011.300308.

Slonje, R., & Smith, P. (2008). Cyberbullying: another main type of bullying? *Scandinavian Journal of Psychology, 49*, 147–154. https://doi.org/10.1111/j.1467-9450.2007.00611.x.

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

Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. *2012 11th International Conference on Machine Learning and Applications (Icmla 2012),* (Vol 2, pp. 386–393). https://doi.org/10.1109/icmla.2012.218.

Tenore, M. J. (2013). Why journalists have trouble with bullying. Retrieved from https://www.poynter.org/reporting-editing/2013/why-journalists-have-trouble-with-bullying/. Accessed 3 Jan 2019.

Ttofi, M. M., Farrington, D. P., Losel, F., & Loeber, R. (2011). Do the victims of school bullies tend to become depressed later in life? A systematic review and meta-analysis of longitudinal studies. *Journal of Aggression, Conflict and Peace Research, 3*, 63–73.

van Geel, M., Vedder, P., & Tanilon, J. (2014). Relationship between peer victimization, cyberbullying, and suicide in children and adolescents: a meta-analysis. *JAMA Pediatrics, 168*, 435–442. https://doi.org/10.1001/jamapediatrics.2013.4143.

Wardecker, B. M., Edelstein, R. S., Quas, J. A., Cordon, I. M., & Goodman, G. S. (2017). Emotion language in trauma narratives is associated with better psychological adjustment among survivors of childhood sexual abuse. *Journal of Language and Social Psychology, 36*, 628–653. https://doi.org/10.1177/0261927X17706940.

Wolke, D., & Lereya, S. T. (2015). Long-term effects of bullying. *Archives of Disease in Childhood, 100*, 879–885. https://doi.org/10.1136/archdischild-2014-306667.

www.nobullying.com. (2016). The top six unforgettable cyberbullying cases ever. Retrieved from https://nobullying.com/six-unforgettable-cyber-bullying-cases/. Accessed 3 Jan 2019.

Young, R., Subramanian, R., Miles, S., Hinnant, A., & Andsager, J. L. (2017). Social representation of cyberbullying and adolescent suicide: a mixed-method analysis of news stories. *Health Communication, 32*, 1082–1092. https://doi.org/10.1080/10410236.2016.1214214.