Correlating Emotional Content in Journalists' Personal and Professional Writing

Michael Grauwde
Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Niels Bohrweg 1, Leiden, The Netherlands
m.n.j.grauwde@umail.leidenuniv.nl

Thesis for the Media Technology MSc Program
Leiden University
June 24, 2022
Thesis advisors: Dr. Maarten H. Lamers & Dr. Suzan Verberne

Abstract

Emotion detection and sentiment analysis are fields of natural language processing that are gaining steam in artificial intelligence (AI). With the emergence of large language models such as GPT-3, more questions are being asked about what can be expected from AI in the future. This study uses text analysis tools to detect emotion in journalists' writing to explore questions of objectivity. Journalism has long been a field that prides itself on being the gatekeeper of truth. However, along with the growth of the internet, there has been a decline in traditional journalism. There has been a rise in opinion pieces and an embracing of emotion in journalism. Twitter has seen journalists interact with their audience on an almost daily basis. This paper uses Twitter as a proxy for personal writing (tweets) and correlates the emotional content in tweets and journalistic articles. By using LIWC – a text analysis programme – this study analyses the emotion (negative and positive) in tweets and journalistic articles and the correlation between the emotional content. Results suggest a significant positive correlation between emotion variables (affect, negative emotion, emotion) in the tweets and professional articles. Even so, this study is limited and requires future research and hopes to function as a foundation for future research in the field.

1. Introduction

In 2021, as the pandemic was in full swing in the Netherlands, I remember reading an article by Dartmouth University professor Bruce Sacerdote called, "Why is all Covid-19 News Bad News?". The article was asking the same question that I was asking myself, though I often asked myself this question about news in general. Sacerdote's query came from the proliferation of prominent mainstream media outlets that, in the year 2020, were covering all Covid-19 news as negative even though there were many vaccines in development and we were in the midst of one of the greatest scientific achievements of the 21st century (Sacerdote, 2020). So, in all the guff, there was some positive news to be shared. I had become fascinated with the reason why we were attracted to such news and, as a follow up, why the news was written in such a manner. It felt to me like asking the question of what came first, the chicken or the egg? Did negative news come first, or did our reaction to news make companies present more negative news?

News and journalism used to be two of the gatekeepers and protectors of truth in our world and allow for the democratic dissemination of information (Waisbord, 2018). Scholar Silvio Waisbord has argued that journalism was among one of the institutions that were closely associated with elite individuals and based its dominance and gatekeeping power on scientific realism (Waisbord, 2018). However, journalism no longer has the aura of being the gatekeeper of truth for many (Waisbord, 2018). As the internet has risen in prominence, social media has as well, and we have seen the decline of traditional journalism, of experts, the rise in partisan-divided trust in facts, and the politicization of science (Baughman, 2015; Dahlgren, 2018). An example of this decline in traditional journalism is the decline of journalists working in outlets and the number of newspapers in circulation (Bird, 2009). Since Bird wrote that article in 2009, news and traditional journalism have continued to decline as social media websites have risen, and many people now get their information directly from social media. While many have also lamented the decline in local newspapers and their impact and circulation (Nielsen, 2015), the internet has brought the dismantling of the knowledge production structure and dissemination that was in position and has been replaced by an information disorder (Waisbord, 2018). The digital environments that have risen up have created different opportunities for public expression with different manners in the way that they engage with news and information. However, these sites like Twitter and Facebook have allowed for a raft of information disquiet and disorder, leading to a raft of what we now refer to as fake news to emerge (Waisbord, 2018).

Emotion has always been part of the journalistic landscape and the profession. However, as Peters (2011) states, the emotion that used to be reflected in journalism used to be that of a journalist being cool. A news report has to be able to catch your attention, be intriguing, uplifting, or scary to get its story across (Beckett & Deuze, 2016). There was always an aversion to having emotions as journalism tried to uphold scientific principles to determine what was the truth (Waisbord, 2018). There has been an increase in the diversity of emotions that is allowed to exist in journalism (Peters, 2011). Mainstream journalism and journalists have sustained their efforts to re-establish themselves as the authority with regard to truth among the rise of fake news in the world (Waisbord, 2018). Emotion has long been a threat to the ideals of journalism, but with the rise of citizen journalism and wider participation in the media landscape, journalism has become more impacted by emotion (Wahl-Jorgensen, 2020). At the same time, since the 1990s, with advances in neuroscience, ideas of rationality and decision-making being unimpacted by emotion have also begun to fall to the wayside (Davis, 2018). Furthering the movement and decline of traditional journalism to birth a new type of journalism.

There has been research on Twitter and other microblogs in relation to analysing the mood of the public. Researchers Bollen et al. have used tweets to find microscopic representations of mood (Bollen, Mao, & Pepe, 2011). They used Twitter because they found it to be quite flexible in its content about what is discussed. They mention that through studying public mood, they find that an individual tweet is, in their words, “a microscopic, temporally-authentic instantiation of sentiment” (Bollen, Pepe & Mao, 2009). Thus, as such, a tweet by an individual can reveal possible inner emotions. Researchers have also used Twitter to see if they could predict the onset of depression in tweets (De Choudhury, Gamon, Counts, & Horvitz, 2013). The reason why social media is used as a paragon of information is for its ability to allow users to share their thoughts and mood with their network, which is revealed by the emotion and language that exists in the posts (De Choudhury, Gamon, Counts, & Horvitz, 2013).

While we are aware that tweets can reveal information about people, we do not know yet if there is a correlation between the moods of journalists, the emotion expressed in their tweets, and the emotion that may exist in their articles. This is important because it opens a dialogue on the impact that social media, particularly Twitter, may have on journalistic output. And furthermore, we do not know how long this impact of emotion may last. For instance, Bollen, Mao, and Pepe (2011) found that the 2008 economic crisis heavily impacted the public’s mood by examining tweets on Twitter in the days during the crisis.

Journalists have used Twitter as a very important tool due to how easy it is to use but also how wide its reach can be. It has effectively changed the way journalists work. Twitter has become a place where journalists, like everyone else, have a network and the ability to communicate within that network (Hermida, 2016). It is also used by people as a place for information dissemination and for getting information about stories instead. Twitter has created a direct line for journalists to reach their audiences. As was seen in research from Pantti (2019) and Oluasson (2017), journalists share their opinions, personal lives, and moral stances on Twitter. The internet has brought the rise of digital journalists.

Other research has explored the wider context of emotion and social media and the place that emotions occupy on social media; this paper will look specifically at journalists. The research question is whether there exists a correlation between the valence of emotional content expressed in journalists’ tweets and their articles in a given time distance? Twitter is a major platform for human expression and allows people to share their ideas with the online world. Twitter can lead to the sharing of objective content, professional writing but also personal content (De Choudhury, Counts & Gamon, 2012). It will analyse the correlation between journalists’ personal and professional writing. Using Twitter as a proxy for journalists’ personal writing and correlating the emotional content found in their tweets with the emotional content found in their professional pieces.

A reason for analysing if emotional content in personal writing on Twitter makes it into professional writing by journalists is to explore the question of objectivity within journalism. Researchers Lee and Hamilton (2022) explored how cognitive bias impacts the word choices of journalists. They analysed news pieces, tweets as well as broadcasts generated in a one year period by 73 campaign reporters during the U.S. presidential election in 2016 using LIWC (Lee & Hamilton, 2022). They explored cognitive biases in journalists' word choices by analysing if journalists engaged in system 1 thinking instead of system 2 thinking. System 1 thinking is a fast, low in effort, and the instinctive manner of thinking that can be centred on emotions and habits. Whereas system 2 is focused more on a slow and analytical manner of processing information to find a conclusion (Lee & Hamilton, 2022). They found that system 1 thinking is fundamental in journalism as there is a focus on social media. Here, we explore the position of objective journalists that may find themselves personally influenced by either certain biases also by happenings in their daily lives. Objective journalists are often journalists who work in serious or traditional journalism focused on the public as opposed to tabloid journalism, which is focused on human interest and personal pieces using less analysis (Pantti, 2010; Skovsgaard, 2014).

To answer our research question, we consider output from a sample of journalists who have written both journalistic and Twitter output at a fixed time distance from each other. With valence of emotional content, we mean its negative, neutral or positive character. Valence is assessed via automated methods. Statistical testing will affirm if emotional content is correlated. When positively correlated, there is a positive correlation between the emotion prevalent in the tweets and the emotion prevalent in the articles or vice versa. When negatively correlated, however, it means that there is an opposite effect. That the emotion prevalent in tweets goes in an opposite direction to the articles and vice versa. When there is no correlation between the two, the two outputs are simply not correlated.

We will collect data for 6 journalists' tweets and articles over a 6-month period, similar to the study by Bollen et al., (2011) which also looked at a 6-month period, and study by Sacerdote et al., (2021) that was a 7 month period. To choose our journalists, we took inspiration from both Pantti (2019) and Fahmy et al., (2022). Pantti's study analysed 3 journalists while Fahmy et al., (2022) selected 13 journalists from 4 different countries. We decided that for our study, a selection of 6 journalists from 3 different organisations could give us some insight and provide results. Then, we will analyse this data using Linguistic Inquiry and Word Count (LIWC). LIWC is a computerised text analysis programme that uses words to find the psychological meanings behind words (Tausczik & Pennebaker, 2010). We used LIWC because it has been used in thousands of studies on text analysis. As people use words to express their thoughts, how they are feeling, and what they are looking out for, LIWC analyses these words to give researchers insights into what those words may mean (Tausczik & Pennebaker, 2010). Emotions can show us how people experience the world, with regards to the valence of the emotion, their expression of that emotion, and the degree of their expression of emotion (Tausczik & Pennebaker, 2010). Research has found that LIWC precisely identifies how emotion is used in language (Tausczik & Pennebaker, 2010). The LIWC 2022 dictionary gives 102 categories and which counts how many times words have appeared in a text file. It uses this bag-of-words approach to show 102 scores, for which each shows the percentage of times that word shows up in the text (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022).

However, LIWC is not free and has academic and commercial licenses. An academic license is € 66,95 a year, while a commercial license can be retrieved after discussing it with the company for researchers attempting to do similar research. The paper by Lee and Hamilton (2022), also using LIWC, helped provide a background for the hypotheses as they also researched positive and negative emotion in their research. Found that journalists used more emotions in tweets than in articles.

We have outlined the hypotheses related to this study:

H_0: *There are no significant correlations between the emotional output in the journalists' tweets and the journalists' articles.*

H_1: *There are significant correlations between the emotional output in the journalists' tweets and the journalists' articles.*

In this way, we will explore how journalists can possibly also be influenced by emotions and how these emotions can overlap from their personal lives and personal writing (on Twitter) to their professional writing (articles).

One motivation behind this study is that we can reveal an important part of how journalists can be influenced by emotions. We also prompt exploration into the role that emotion has in journalism in what we now know as a post-truth landscape. Furthermore, a larger question on emotion and intelligence can be encouraged as we explore how an occupation supposed to be synonymous with objectivity and reason can, in fact, be influenced by emotion.

We analysed tweets by 6 journalists over a 6 month period. Then correlated their concatenated tweets for each professional piece of journalism. In the end, we had a selection of tweets for the week prior to the day the article was published. In this manner, we matched the article with a selection of tweets from one week prior to the article's publication. This allows us to analyse the emotional content of journalists' personal and professional writing. This question is delicate, however, as we must make sure not to infer broad conclusions on the basis of technology that still examines an output of an individual in their text and does not ask them questions on a personal level. These programmes currently focus on the explicit text of writers rather than focusing on the implicit. An example here comes from research written by Van Hee, Lefever, and Hoste (2018), where they reinforce the lack of computers' ability to detect segments of language that humans can detect, such as irony. There is a lack of a wider knowledge of the world that exists that restricts current tools from understanding this segment of language. So, while Linguistic Inquiry and Word Count (LIWC) provides a wide breakdown of a text, it can fail in its inability to receive the context of the wider story. However, it is possible that text analysis can add another tool to the tool belt when asking broader questions about our societies.

2. Background and prior work

2.1. Journalism and Emotion in a New Landscape – A changing landscape

In the years since World War II, journalism landscapes in both Britain and the United States of America were very different (Baughman, 2015). While the USA had three times the population of

Britain and also far more newspapers in circulation, the British population was still far more likely than American citizens to read a newspaper every day (Baughman, 2015).

After the war, newspapers also started to become less deferential towards the government. This occurrence coincided with the decline of a large number of newspapers being owned by one person or by corporations. And so, the decline in traditional journalism began (Baughman, 2015). While journalism was seen as the Fourth Estate, as it became more independent and critical of the government, the first attack on traditional journalism came from opinionated news programming (Baughman, 2015). This attack was further exacerbated by the rise of the internet. As the internet presented a large threat to daily newspapers, many large newspapers were unable to see the impact that the technology would have on their operations (Baughman, 2015). The demand for on-demand news and, in particular personalised news and journalism grew as the internet became cheaper to access. In America, the number of newspaper editorial employees fell from 60,000 employees in 1992 to 38,000 in 2012 (Baughman, 2015). While newspapers like the Guardian, previously known as the Manchester Guardian and a longstanding destination for serious and traditional journalism, and the New York Times have survived, there is a clear distinction in the way that they now present the news (Baughman, 2015).

Journalism is often seen as being objective and almost scientific in the manner that it goes about collecting correct information to report to the larger population. The institution of journalism in the latter half of the 20th century was one that considered itself the gatekeeper of truth and news (Waisbord, 2018). However, from the start of the 21st century, a swift change began to occur in the world of journalism. As the world became more digital, journalism changed, and often, there are multiple perspectives by which truth is ultimately determined (Waisbord, 2018).

At the beginning of the 1990s, advances in neuroscience found that emotions were vital in decision-making (Davies, 2018). This can be due to emotion-inducing stress or leading to an incidental affect on the ability of people to make decisions (Peters, Västfjäll, Gärling & Slovic, 2006). This research can also be supported by George and Dane (2016), that investigated effects such as the disposition effect, finding that in finance, traders often sell stocks that are winning too early and often hold on to stocks that are losing too long. Here regret and elation were seen to be connected to these two actions. The researchers also found that how people are currently experiencing life should impact how they make decisions (George & Dane, 2016). Furthermore, affect needs to also be understood to see how team members’ decisions are influenced (George & Dane, 2016).

As journalism becomes more emotional due to the changing media landscape, where being emotional used to be related to irrationality or lacking objectivity, being emotional is now essential in the way that news is presented (Davies, 2018). Emotions are integral to our decision-making process, and, as such, are an essential by-product of journalism. However, due to a change in sensitivity, emotions can lead to disruptive situations if they are not treated with care in the media landscape (Davies, 2018).

Emotion is not a new thing in journalism. Emotion has always existed at the centre of journalism to keep the audience engaged with the news stories being reported. At the start of his article, “Emotion aside or emotional side? Crafting an ‘experience of involvement’ in the news”, Chris Peters (2011) shows us various examples when journalists were influenced by emotion due to the emotional things that were happening at the time, or rather also stated a need for information dissemination to be based on some semblance of emotion (Peters, 2011). He shows us the position of Walter Cronkite and how he struggled to remain himself following the assassination of President Kennedy (Peters, 2011). Peters points out that it is not that journalism itself has become more emotional, as it always was, as he gave an example of Walter Cronkite, but furthermore that there has emerged a diversity in the type of emotion in journalism as well as the attempts to involve the audience (Peters, 2011). It is not that emotion is new to us. It is that it has become embraced by some semblance of our society that yearns for emotional content in the way that news is presented.

Massive changes in technology and the digital landscape have allowed emotion to become a more important part of the dissemination and production of news stories (Orgeret, 2020). With the rise of technology and wider digital sharing on social media, we have seen the rise of what is known as “citizen” or “amateur” journalists (Wahl-Jorgensen, 2020). This has allowed the emergence of more first-person storytelling that lacks the objectivity that often exists in professional journalism. Orgeret mentions that journalists’ work these days is often subject to commenting and resharing via Twitter with the larger public (Orgeret, 2020). Beckett and Deuze have stated that emphasizing emotion as a key theme in journalism redefines objectivity in journalism and, in fact, reshapes news itself (Beckett & Deuze, 2016). Beckett (2015) states that with technology and social media came the advent of “networked news”. News is networked as news is ubiquitous and impossible to escape. We are in a sea of information which has meant that news is more difficult to digest (Beckett & Deuze, 2016). In this manner, emotion has become a key aspect of producing news for the world. Emotion has become a key theme used by journalists due to economic, technological, and behavioural factors (Beckett & Deuze, 2016). The rise of competition in the media space has led to changes in how news is presented and the styles in which articles are written. Social media has become vital as outlets attempt to involve their audience (Beckett & Deuze, 2016). Technological factors have provided us with data that shows that emotional cues help to get the attention of consumers and their engagement, and journalists have also amended the style in which they write their articles to pull in a larger audience (Beckett & Deuze, 2016). And on the behavioural factor, now, there is a wider idea of understanding why people behave a certain way. From this information, we see that people respond to emotion rather than facts, and in this way, the media has endeavoured to sell consumers news (Beckett & Deuze, 2016).

Beckett indicates that with social media and with their networks, there are interesting feedback loops between journalists and the responses that they receive from their network (Beckett, 2015). The emotion that exists in connection with the journalists’ work can have an impact on the way they report the news (Beckett, 2015). Pure objectivity, which has always been an aspiration of journalists and journalism is beginning to take a backseat to emotion, and this can lead to filter bubbles in that we only want to hear ideas and views that align with our emotional sentiment (Beckett, 2015).

Scholar Mervi Pantti states that emotional journalism used to be associated with journalism that was more popular and tabloid and focused on everyday life and thus focused on emotions (Pantti, 2010). Tabloid journalism is a style that is more focused on narratives and limits its use of analysis while focusing more on personal and human interest articles. Researchers also point out that sensationalism is vital in tabloids (Skovsgaard, 2014). In contrast, serious “quality” journalism was focused on information and the wider public (Pantti, 2010). News was a place for discussion driven by objectivity, according to journalism. And that change due to possible market forces has, according to some, led to an increase in emotion in journalism and to decay of quality journalism (Pantti, 2010). When it comes to journalism in emotion, journalists have argued that emotion helps to get the story across. Journalists argued in the paper by Pantti that emotion shapes how viewers watch and, in a way, how they engage with the narrative content (Pantti, 2010). The question journalists often ask is why to get emotion across, is it for commercial profit and better ratings, or is it the best way to tell the story (Pantti, 2010). Journalists in the paper by Pantti stress that emotion does not change the traditional values of objectivity within journalism. Rather, it is more focused on emotional storytelling to provide a better story to the audience. In a 2019 paper, Pantti addresses the position that Twitter itself has taken in the dissemination of visual images within conflict reporting regarding the Ukraine conflict with Russia since 2014 (Pantti, 2019). In that way, the emotional and personalised level of reporting is further examined. Emotion can find its way into these visual images that came out of the conflict at the time (Pantti, 2019). Furthermore, there has been research done on the use of Twitter by journalists. Exploring the balancing act between remaining an objective information provider and expressing emotions. This research has found that journalists that are more enthusiastic users of Twitter are more likely to challenge the position of objectivity and neutrality within journalism.

2.2 Twitter and Journalism

Since Twitter's founding in March 2006, the media landscape has been transformed and has allowed journalists to reach their audience easily, with 75 percent of UK journalists using Twitter in 2015 (Hermida, 2016). Research has found that Spanish journalists are quite open when measured in participatory terms with respect to the high level of direct replies that they have on Twitter (Noguera Vivo, 2013). Moreover, users also tend to associate Twitter as a source of news more than Facebook (Hermida, 2016). Twitter has also allowed journalists to gather news from a variety of sources in real-time (Hermida, 2016). The emergence of Twitter has blended the professional and personal lives of journalists as journalists mix their personal experiences on Twitter and, as such, has blurred the lines (Hermida, 2016). Including using Twitter as a way to reach their audience, Twitter is also important for journalists to create a personal brand – something that is often recommended – this is seen more predominantly among journalists who use Twitter often (Hedman & Djerf-Pierre, 2016). However, many journalists are critical of personal branding on Twitter and deem it incompatible with journalistic norms. Furthermore, while those journalists who blur the lines between personal and professional have a positive outlook toward audience interaction, this personalisation, which often leads to more transparency, also leads to more online threats and critique (Hedman & Djerf-Pierre. 2016).

Research from Canter (2015) found that local journalists in Bournemouth at the Bournemouth Daily Echo (a small local newspaper) were not big users of Twitter for personal information and used it mostly for their jobs, even though they engage in personal branding. Olausson (2017) found that journalists that were heavily followed on Twitter did use their platform to engage in personal interactions with their audience. Personalised reporting is a term that was developed by Mervi Pantti in her 2019 study on personalised reporting in the Russian invasion of Crimea. However, a similar term was used by Canter, *personalised tweeting*. This personalisation of Twitter has become normalised as Twitter has developed into a necessity for journalists. Journalists who use Twitter are shown to be more personal in the tweets that they share online (Lasorsa, Lewis & Holton, 2012). Conflict reporting is a subsection of journalism that can provoke journalists into assigning their own feelings and judgments towards the conflict. Twitter can provide an outlet for these feelings.

In researching the war in Ukraine in 2014, Pantti showed that the imagery that came out of the war showing the plight of Ukrainians living through a prolonged conflict in the Eastern regions was not a focus (Pantti, 2019). Journalists often act as the visual gatekeepers on Twitter by disseminating the visual content that emerges from the war (Pantti, 2019). Journalists that Pantti studied found themselves attributing their personal feelings and judgments to their tweets while reporting on Twitter (Pantti, 2019).

In Pantti's study, she found that there was a combination of fact-based neutral reporting and the sharing of personal stories and opinions (Pantti, 2019). There were image tweets shared by journalists that focused on the sharing of breaking news, such as the crash of airline flight MH17, which often was shared through retweets or eyewitness photos. However, there was another level of image tweets that were more subjective and reflected the journalists' moral claims, emotions, and opinions that in traditional media would need to be separate from reporting (Pantti, 2019). Here, Pantti argues that the journalists themselves became protagonists in the information war, and journalists were pulled into moral questions of what is right and wrong (Pantti, 2019). Here, Pantti states that the discourse of the image tweets was limited to upholding the perspective of Western media regarding the Ukraine conflict while ridiculing Russian authorities or media (Pantti, 2019). She states that the tweets' emotional power was not used to focus on the humanitarian narrative of the suffering of civilians (Pantti, 2019). She also states that the focus of the tweets was on the journalists themselves rather than the news topics themselves (Pantti, 2019). While there is nothing new about journalists' position as neutral onlookers or moral agents, the combination of these roles within the world of journalism becomes a part of the narrative told on Twitter. The information war on Twitter, according to Pantti had a clear impact on the emotional discourses that the image tweets followed (Pantti, 2019).

Research on the Yemen civil war found that Yemeni journalists focused mostly on the aftermath of the war, which did indeed suggest emotional tones (Fahmy, Taha & Karademir, 2022). However, as a whole, the research found that the reporters mostly were neutral in tone. However, there was a difference in the tweets that were shared (Fahmy, Taha & Karademir, 2022). Yemeni journalists focused on everyday life and violence in Yemen, while non-Yemeni journalists focused on the larger international politics and the consequences of the armed conflict (Fahmy, Taha & Karademir, 2022). Thus, there were tweets that were subjective in nature from the many journalists reporting on the conflict (Fahmy, Taha & Karademir, 2022). As such, the research shows that Twitter can lead to personalised reporting and subjectivity in what is shared with the journalists’ audience. Questions of truth in journalism can then arise.

2.3. Cognitive Bias and its impact on Journalism

Social media and emotion on social media by journalists can make journalism more personalised, and so there is a possibility that we can be inundated with personalised propaganda (Beckett & Deuze, 2016). These can be further exacerbated by the feedback loops mentioned earlier (Beckett, 2015). These tweets can lead to fundamental cognitive biases such as motivated reasoning and confirmation bias. Motivated reasoning is the tendency in people to only notice information if it aligns with their outlook. This is something that, in the case of echo chambers on social media, can easily occur as we only see information within our bubble and are far more critical of information that may pierce that bubble or may disregard it completely (Kahan, 2013). In this way, we conform with our communities that already agree with our views and become more ideological or biased in our worldview (Kahan, 2013). When we become emotional, we become less rational and sound in our decision-making and become less interested in what may challenge our beliefs.

In journalism, as in daily life, cognitive biases persist in many areas, such as decision-making, stereotyping, and assigning causes to certain events (Christian, 2013). Stereotypes, even in a subconscious manner, have the ability to influence thinking as individuals tend to focus on certain information as well as interpreting information in such a manner to match the stereotype. In this way, they serve as hypotheses for comprehending new information in a disproportionate manner (Christian, 2013). While there is a substantial number of research that has been done in the field of stereotyping in media, this is only one cognitive bias that is analysed. While studying a specific case in the United States of a crime at Duke University in part of the lacrosse team, researchers found that the media made caricatures of the individuals while also oversimplifying and misrepresenting what the case was about (Entman & Gross, 2008). This case caused a wider discussion in the United States of America around questions of race and media bias (Entman & Gross, 2008). An analysis of the United States of America media’s representation of the Iraq War in the periods of May 5th to May 26th, 1998, October

11th to the 31st, 2002, and May 1st to May 21st, 2003 found that there was a parroting of the White House's perspective, in particular on the existence of weapons of mass destruction in Iraq (Moeller, 2004). There was also a lack of an alternative perspective to the line presented by the White House (Moeller, 2004). The study also found that this representation is prevalent in the way that the media presented other international stories, with a lack of independent analysis (Moeller, 2004).

For this reason, it is intriguing to look at the way that journalists may frame their tweets and the overlap that the emotion in their tweets may have with their articles to examine if there is some level of cognitive and ideological bias deluged in their thinking. It would allow journalists to become more aware of their emotions at the time of their writing and to further reflect on their cognitive biases to make for more objective journalism.

2.4 Sentiment Analysis and Emotion Detection

Humans have long used language to communicate messages, and we use text in combination with our language to convey different emotions and ideas (Kaur & Saini, 2014). Using computers, we are able to classify these texts according to emotions (Kaur & Saini, 2014). Sentiment analysis, also known as opinion mining is a computerised method to find the opinions of authors about certain things (Feldman, 2013). Sentiment analysis allows companies and researchers to use social media sites such as Twitter to gain information about the mindset of their target audience (Feldman, 2013). Sentiment analysis often classifies text as either negative or positive, and while it is the overlying method associated with opinion mining, when attempting to detect emotions in text, there is a slight difference (Kaur & Saini, 2014). Emotion detection dives deeper into dividing text between a variety of emotional states and allows the user of the classification to differentiate between emotions such as happy, sad, anger, and disgust (Kaur & Saini, 2014). In this article, we will be focusing squarely on emotion detection as a subfield of sentiment analysis.

Emotion detection allows us to receive more information than the simple negative or positive that sentiment analysis allows us. However, in turn, this makes the problem of using classifiers more difficult (Taboada, 2016). Machine learning approaches are able to collect data based on prior information that has been collected, but difficulties emerge when collecting data based on specificities such as emotion. An approach often used in sentiment analysis is the lexicon-based approach. This approach allows sentiment values to be derived from certain words in the text analysed based on the lexicon used (Taboada, 2016). An interesting exploration of the opinions conveyed online shows us that there is a higher frequency of positive words being used online than negative. However, there remains a huge negativity bias where negative words and events have an outsized impact on our psychological state and behaviour (Taboada, 2016). Emotion detection allows us to dig deeper into these emotions and how they may have an impact on our decision-making. Either positive or negative. As such, this article will be focusing squarely on emotion detection as a subfield of sentiment analysis.

2.5. Emotion Detection

Emotion detection allows us to explore emotions such as fear and anger, which are both negative emotions to be distinguished from each other (Seyditabari, Tabari & Zadrozny, 2018). This has exploded in research in recent years in the fields of psychology, marketing, and computer science (Seyditabari, Tabari & Zadrozny, 2018). Seyditabari, Tabari, and Zadrozny (2018) mention that emotion detection systems can be used to analyse the emotions in authors' texts. Understanding the role of emotion in decision-making can allow companies to determine how people respond to products and services. They also suggest that it can allow for the creation of better AI tools such as chatbots (Seyditabari, Tabari & Zadrozny, 2018).

The authors also mention that the complexity of our emotional state, as well as emotional context and usage of metaphors used in text, are among some of the problems that emotion detection comes up against (Seyditabari, Tabari & Zadrozny, 2018).

Many of the current emotion detection systems are based on research by Ekman on basic emotions (Ekman, 1999). These emotions are "anger, surprise, disgust, fear, happiness and sadness" (Ekman, 1999). Other researchers have added more emotions, such as shame and pity, to the list of the six mentioned by Ekman, arguing that emotions differ in intensity (Acheampong, Nunoo-Mensah, & Chen, 2021). The Circumplex of Affect model has also been used in connection with emotion detection models. This model deals with valence, arousal, and dominance and is used together with transformer (deep learning) models to gain further information on emotion from texts (Acheampong, Nunoo-Mensah, & Chen, 2021).

As we mentioned earlier, the field of emotion detection is quite a new field; however, there are many researchers that are already well versed in the field. Choudhury, Gamon, Scotts, and Horvitz have used emotion detection in tweets and have found out how to recognise severe depression via tweets (2013). As we mentioned earlier, Bollen, Mao, and Pepe (2011) researched if tweets could represent the wider public mood using emotion detection, while Bollen, Mao, and Zeng (2011) also researched if Twitter could be used to predict the stock market. In their analysis, they found that the public mood on Twitter was around 85% accurate in predicting the movements of the Dow Jones Industrial Average (Bollen, Mao & Zeng, 2011). On the more applied research side, Hasan, Rundensteiner, and Agu (2019) developed Emotex, a system based on the Circumplex Model of Affect that detects the emotion of tweets during live streams of text.

Research by such as Chatterjee et al. (2019) and Hasan et al. (2019) used methods of classification based on emoticons and hashtags, respectively while exploring a larger number of tweets. Our study will take a different approach to analyse the tweets of journalists as well as their articles.

2.6. The Growth of Text Analysis

My interest in text analysis and what it could reveal to people was sparked by research from scholar Ian Lancashire and computer scientist Graeme Hirst. They initially researched Agatha Christie specifically before researching Agatha Christie’s work in tandem with the work of Iris Murdoch and P.D. James. Agatha Christie was suspected of having Alzheimer’s while Murdoch died with it, as P.D. James died healthily (Lancashire & Hirst, 2009; Le, Lancashire, Hirst & Jokel, 2011). This research allowed them a peek into the development of the minds of these novelists and examined if they would be able to detect dementia by analysing text through time (Lancashire & Hirst, 2009; Le, Lancashire, Hirst & Jokel, 2011). The research found that it was possible that Agatha Christie did suffer from Alzheimer’s while she was working on her last novels and that the same was true of Iris Murdoch, who possibly suffered from dementia (Lancashire & Hirst, 2009; Le, Lancashire, Hirst & Jokel, 2011).

This interest led me to work done in text analysis by using LIWC. LIWC is a computerised text analysis programme that was developed to analyse text and find the psychological meaning of words (Tausczik & Pennebaker, 2010). It has been used to analyse emotion, such as in depression detection, while it has also been used in analysing short texts and in social media research. LIWC has been used in predicting depression by analysing tweets and has also been used to spot trends in stress also by studying tweets (Shen et al., 2017; Wang, Hernandez, Newman, He & Bian, 2016). As one of the most widely used computerised text analysis methods that analyses emotion by way of positive and negative emotion, we decided to use it in our study (Howes, Purver & McCabe, 2014).

3. Material & Methods

The purpose of this study is to analyse whether there is a correlation between journalists’ personal writing (tweets) and professional writing (articles) with regard to emotional content. To analyse this data, we planned to collect tweets and articles from 6 different journalists over a 6 month period. For each selected journalistic article, we collected the tweets of the same author from the week prior to the article's publication date. In this way, we have multiple data points that are categorised by the date of the article’s publication.

While the length of time is important, more important to us is analysing these journalists’ emotional content on their Twitter and in their articles. More specifically, we will be analysing whether there is any correlation between the two.

3.1. Journalist Selection

For our purposes, journalists need to use Twitter and need to be writing quite often for their news organisation. As a result, we are looking at news organisations that are quite widespread in their reporting, reporting both in their home country and with a vast presence abroad, so we could have news organisations that are widely read. In this manner, we could find influential and reputable news organisations. We also wanted to analyse news organisations that, if they are biased in any way, are all over the political spectrum. So, it was vital for us to find news organisations that were politically on the left, centre, and on the right. Also important to us was finding news organisations that are quite reputable in their fact-based reporting but also for their analysis-focused articles. This is so that they meet the criteria of serious journalism and that the journalists analysed are serious journalists. Ad Fontes media is a company that analyses the reliability and bias of news organisations. They found that all three organisations in our case are reputable and report fact-based reporting as well as analysis (Ad Fontes Media, 2022). Another essential aspect for us was to find news organisations available in the same language, in our case, English, and to find news organisations that were widely read. As the research takes inspiration from Sacerdote's research on news from the United States of America, we decided to focus on English news organisations also. Another reason for this focus is because the region of news interesting to me is English. Lastly, we also wanted to be sure that we could find methods that could analyse news in the language. Thus, we decided to select journalists that report on world news or analysis for the New York Times, Reuters, and the Wall Street Journal. The choice of the news organisations is based on their possible different political biases on the political spectrum (Ad Fontes Media, 2022). The New York Times is seen as slightly to the left, Reuters as more towards the centre, and the Wall Street Journal is slightly to the right (Ad Fontes Media, 2022). Journalists we are looking at the need to have a Twitter account and also have to write often for their news organisations. We also chose organisations that are widely circulated and, as such, widely read. Reuters has a vast readership of 40,574,000, New York Times an audience of 162,890,000, and Wall Street Journal 39,430,000, so these are all widely read news organisations (Ad Fontes Media, 2022). An important requirement for being able to research this topic is to be able to search and find all of the articles available on the website of the news organisation written by the individual journalist.

It is necessary for our research question to analyse journalists that have a high position in their organisation, either as bureau chief or head correspondent for their different regions. This is once again so that we could fit the criteria of analysing serious journalists and serious journalism. So, we attempted to explore these options while also exploring countries that would be interesting to analyse to see if there is this question of emotional correlation. Most of the journalists we considered work in the Asia Pacific region and in Europe. For each news organisation, we selected two journalists. One is more focused on reporting in Europe or a European country, and one is more focused on the Asia Pacific region. We thought that journalists analysing politics around the world might be more susceptible to emotional and ideological bias because politics is an emotionally charged field (Bleiker & Hutchison, 2008).

The journalists that we selected to analyse for this research are:

1. Carlotta Gall. She is a journalist from the New York Times that is currently the Istanbul bureau chief for the New York Times and has been at the organisation since 1999. Though she covers Istanbul, she often also reports on Pakistan and Afghanistan. Her Twitter profile is: **https://twitter.com/carlottagall**

2. Hannah Beech. She is a journalist from the New York Times that is currently the senior correspondent for Asia and is based in Bangkok, Thailand. She covers around 10 countries and previously worked at The Times for around 10 years as the Southeast Asia bureau chief. Her Twitter profile is: **https://twitter.com/hkbeech**
3. Yaroslav Trofimov. He is a journalist from the Wall Street Journal and is the chief foreign-affairs correspondent for the Journal. He covered the Taliban takeover of Afghanistan in 2021 and, in 2022 was working in Ukraine. Trofimov has been at the Wall Street Journal since 1999 and was previously the Rome, Middle East, and Singapore-based Asia correspondent. His Twitter profile is: **https://twitter.com/yarotrof**
4. Saeed Shah. He is a journalist from the Wall Street Journal that often covers Afghanistan and Pakistan. His Twitter profile is: **https://twitter.com/SaeedShah**
5. Poppy McPherson is a journalist from Reuters who acts as a special correspondent. She currently covers Southeast Asia, and she covers Myanmar the most. Her Twitter profile is: **https://twitter.com/poppymcp**
6. Francesco Guarascio. He is a reporter from Reuters and is a European Affairs Reporter and is based in Brussels. He often reports on the happenings of the European Union as a block. His Twitter profile is: **https://twitter.com/fraguarascio**

These journalists were chosen based on how often they report and how often they tweet. This way, we collected a dataset that gives us a run-down of the activity of these journalists, both professional and personal, between the period of July 1, 2021, and December 31, 2021. After collecting a large dataset, we hope to get a somewhat representative reflection of the writing of our journalists and of the emotional correlation, if there is any. In this way, we can show if there is a generalisation that may happen throughout journalism itself.

3.2. Text Analysis

The tweets and articles will be analysed with the LIWC system. LIWC is a computer text analysis programme that analyses text by using social and psychological factors (Tausczik & Pennebaker, 2010). We will use LIWC to examine the emotional states of people by way of their tweets and then by their articles. Words like “happy”, and “excited” are words that will probably show that an individual is feeling happy and positive, and these are words that we can use to analyse the emotional state of our journalists (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). As LIWC only examines specific words in the text, it is possible that it can miss some important factors. Emotion focuses on words like “laughter” that express true emotion labels or words that heavily indicate emotions like “good”, or “happy” (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). Suggesting behaviour that possibly flows from a positive affective state (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). At the same time, sentiment variables are variables that reflect sentiment and not only emotion. These are words like “happy”, “angry”, “sad”, etc., while also including words like “beautiful” and “kill” which are related to those emotion words and reflect sentiment (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). In the case of LIWC dictionaries, the variable tone reflects the sentiment of the words analysed (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022).

3.3. Emotion Extraction

LIWC provides a percentage based on the text that shows how much emotion is in the text. LIWC assigns many variables that can be associated with emotion. These are affect, tone, positive tone, negative tone, overall emotion, positive emotion, and negative emotion factors. Affect is the umbrella variable under which positive tone, negative tone, overall emotion, positive emotion, and negative emotion are all sub-variables. Tone provides a percentage to reflect the sentiment that exists in the text, while emotion focuses squarely on the emotion (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). Tone is a summary variable and is calculated by looking at the difference between the LIWC words for positive emotion and negative emotion (Cohn, Mehl & Pennebaker, 2004; Monzani, Vergani, Pizzoli, Marton & Pravettoni, 2021). Overall emotion, negative emotion, and positive emotion are not totally independent from each other as negative emotion and positive emotion words are also part of the overall emotion variable. The overall emotion term is made up of positive and negative emotion, which are analysed in this study, but also anxiety, anger, and sadness words (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). Furthermore, negative emotions and positive emotion values go from 0% to 100%, showing the percentage of negative and positive emotion words that are in the text (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). Emotion words are compared with the different emotion categories (overall emotion, positive emotion, negative emotion, anger, sadness, anxiety) (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). Based on that, the words are then matched to words if there are fragments in the LIWC dictionary of emotion, which then counts all the emotion words in the emotion dictionary (McDonnell, Owen & Bantum, 2020). Then LIWC uses the results of the word count to show the percentage of the words in the text containing emotion (McDonnell, Owen & Bantum, 2020). For example, if negative emotion is 3.2, then 3.2% of the words are negative emotion words.

In this sense, we will be able to understand how much emotion lies in the text that we are analysing, both in the tweets and in the articles. Other sentiment analysis tools may have an advantage over LIWC in that they analyse the context of the text and the word placement in the text, which is something that LIWC overlooks. This point is something to keep in mind while analysing the data. LIWC has been used, however, to study both long and short forms of text, which is something that other models may have trouble with. For example, LIWC has been used to study depression in tweets, anxiety on Reddit, predicting elections based on political sentiment on Twitter, and was also used to study sentiment in textbooks (Park, McDonald & Cha, 2021; Shen & Rudzicz, 2017; Tumasian, Sprenger, Sandner & Welpe, 2010; Sell & Farreras, 2017). Lee and Hamilton (2022) explored the existence of cognitive biases in word choices by journalists finding that journalists did engage in system 1 thinking are more prone to cognitive errors. They did so by analysing if journalists used more emotions in their tweets than in articles (Lee & Hamilton, 2022).

3.4. Correlating Emotional Data

After gathering all the values received from LIWC, we placed them in an excel file and analysed all of the data points. In analysing the data points with each other, we perform correlations between the variables. We will do these correlation analyses with Pearson's correlations in JASP. Pearson's correlations were also used in another LIWC study where narcissism was analysed to find correlations in variables (Holtzman et al., 2019). This research correlated the narcissism term with all LIWC variables, finding significant correlations with some LIWC categories (Holtzman et al., 2019). Here, we will then see if the results fit our H_1 (alternate hypothesis) and if there is any correlation between professional and personal writing for the dataset.

3.5. Dataset

The dataset was made up of 221 articles and 194 concatenations of tweets in our dataset. A concatenation here is the linking of tweets that were published in the week leading up to the publication of an article. So if an article was published on the 8th of July 2021, and for the period of the 1st of July 2021, up until and including the 8th of July, 2021, there were 8 tweets, those 8 tweets would be grouped together. This allows us to get a full picture of the journalists' emotions leading up to the publication of the article. Here, however, we had some issues as some journalists were more active on Twitter than others. Due to some journalists being less active on Twitter, there are some weeks in which there are no original tweets by that journalist. As a result of this, there are 221 articles that are analysed while the number of concatenations that are analysed is 194. So, 27 articles have no corresponding tweets from the same week. The articles were collected manually, while the majority of the tweets were scraped using a snscrape method. Tweets by journalist Francesco Guarascio could be manually taken from Twitter's API as often he also tweeted in languages other than English.

4. Results

4.1 Descriptive Statistics

The research question is whether there exists a correlation between the valence of emotional content expressed in journalists' tweets and their articles in a given time distance. To answer this question, we performed Pearson's correlations in JASP after analysing all of the concatenations of tweets and articles in LIWC. First, we analysed the descriptive statistics for the relevant variables from the LIWC analysis. We did this for both the journalists' articles and for the journalists' concatenated tweets. For the 221 articles and 194 concatenations of tweets in our dataset, there was a mean of 8.08 tweets per concatenation for one article, while the minimum was 0 and the maximum was 32 tweets per concatenation.

The variables of interest are sub-variables of affect in LIWC, overall emotion, negative emotion, and positive emotion, as well as tone variables (overall tone, positive tone, and negative tone). Specifically, we are interested in these variables: emotion (overall emotion, negative emotion, positive emotion), tone (overall tone, positive tone, negative tone), and affect. These variables will show us the amount of emotion words in both the tweets and articles (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). So these would be words like “laughter”, “good” and “happy” for positive emotion and “bad” and “tired” for negative emotion, and all would be under the term emotion (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022). Tone reflects sentiment as well as these emotion words, so these words would also be included in the tone variable. Tone reflects the sentiment, and there are also words that fit specifically under negative tone and positive tone, like “too much” for negative tone and “new” for positive tone (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022).

The study found a difference in affect for the articles having a mean of 3.59%, and tweets for affect, having a mean of 7.95%. Overall emotion has a mean of 0.54% in the articles, while it has a much higher mean of 3.81% in the tweets. As well, this trend is seen in our variables of concern in positive emotion and negative emotion, with the mean for the articles being 0.15% and 0.34%, respectively while in tweets, there was a mean of 0.41% and 3.37%, respectively. Here, we can already see the predominance that negative emotion has over positive emotion. This trend is similar for the positive tone and negative tone, which for the articles is 1.52% and 2.03%, and for tweets is 2.52% and 5.40%, respectively. The descriptive statistics in this paper supported the study by Lee and Hamilton (2022) that found more use of emotion words and system 1 thinking in tweets than articles. However, tone is higher in the articles with a mean of 18.48% than in the tweets, which has a mean of 14.39%, which goes against the study of Hamilton and Lee (2022).

The standard deviation and ranges for all of the descriptive statistics can be found in Table 1 for articles and table 2 for tweets. For the articles, the was a range of 0.0% and 1.14% for negative emotion showing the small amount of negative emotion words in the articles. The range of the variable overall emotion was 0% to 2.48% in the articles. The amount of tone was very large in the articles as it was from 1.0% to 95.61%. The range of the negative emotion in the tweets was 0.0% to 20%. The range of the variable emotion in the tweets was also 0% and 20%; see Table 2. This for the tweets was from 1% to 99%, which is reflective of the tone variable in LIWC (Boyd, Ashokkumar, Seraj, & Pennebaker, 2022).

Table 1. Descriptive Statistics

Descriptive Statistics of the Journalists' Articles (N=221)

	Author	A - Tone	A - Cognition	A - Affect	A - Positive Tone	A - Negative Tone	A - Overall Emotion	A - Positive Emotion	A - Negative Emotion
Valid	221	221	221	221	221	221	221	221	221
Missing	0	0	0	0	0	0	0	0	0
Mean		18.48%	7.74%	3.59%	1.52%	2.03%	0.54%	0.15%	0.34%
Std. Deviation		14.28%	1.75%	1.40%	0.82%	1.18%	0.41%	0.21%	0.28%
Minimum		1.00%	3.85%	0.23%	0.00%	0.00%	0.00%	0.00%	0.00%
Maximum		95.61%	15.32%	7.01%	6.32%	5.71%	2.48%	1.59%	1.14%

(*A refers to the variables based on articles, whereas T refers to variables from tweets)

Table 2. Descriptive Statistics

Descriptive Statistics

	Author	Number of Tweets	T - Tone	T - Cognition	T - Affect	T - Positive Tone	T - Negative Tone	T - Overall Emotion	T - Positive Emotion	T - Negative Emotion
Valid	221	221	194	194	194	194	194	194	194	194
Missing	0	0	27	27	27	27	27	27	27	27
Mean		8.08	14.39%	7.82%	7.95%	2.52%	5.40%	3.81%	0.41%	3.37%
Std. Deviation		8.15	23.91%	4.11%	4.30%	2.13%	4.29%	2.88%	0.78%	2.85%
Minimum		0.00	1.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Maximum		32.00	99.00%	20.00%	30.00%	10.20%	30.00%	20.00%	7.14%	20.00%

(*A refers to the variables based on articles, whereas T refers to variables from tweets)

4.2 Correlation of Emotion Variables

From the correlation matrix below, we analysed the correlation among our variables of interest. These variables are of interest for this study because the correlation between the tweets and the articles with each variable will show us that emotional content is correlated between the articles and the tweets. For positive emotion in articles and tweets being positively correlated, that would show that as positive emotion in articles will go up, so will the positive emotion in tweets. If negative emotion in tweets and articles is negatively correlated, then as the articles increase in their negative emotion, the tweets will decrease in their negative emotion. Pearson's r for the correlation analysis shows if a correlation is weak (<0.1), small (0.1 – 0.3), moderate (0.3-0.5), or large (>0.5).

Results from Pearson's correlation analysis can be found in Table 3. A significant positive correlation was found between negative emotion in the articles and negative emotion in the tweets (r=0.43, $p < 0.001$) (**H$_1$**), but not for positive emotion in the articles and the tweets (r=0.07, n.s.), which was an interesting result.

Table 3. Correlation Matrix for Emotion Variables

Pearson's Correlations for the Articles and Tweets written during the 6-month period									
Variable		**A - Affect**	**A - Overall Emotion**	**A - Positive Emotion**	**A - Negative Emotion**	**T - Affect**	**T - Overall Emotion**	**T - Positive Emotion**	**T - Negative Emotion**
1. A - Affect	Pearson's r	—							
2. A - Overall Emotion	Pearson's r	0.53 ***	—						
3. A - Positive Emotion	Pearson's r	0.32 ***	0.69 ***	—					
4. A - Negative Emotion	Pearson's r	0.50 ***	0.82 ***	0.20 **	—				
5. T - Affect	Pearson's r	0.21 **	0.26 ***	0.18 *	0.18 *	—			
6. T - Overall Emotion	Pearson's r	0.41 ***	0.46 ***	0.29 ***	0.38 ***	0.71 ***	—		
7. T - Positive Emotion	Pearson's r	0.05	-0.07	0.07	-0.14	0.22 **	0.18 *	—	
8. T - Negative Emotion	Pearson's r	0.39 ***	0.49 ***	0.27 ***	0.43 ***	0.66 ***	0.96 ***	-0.10	—

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

(*A refers to the variables based on articles, whereas T refers to variables from tweets)

Furthermore, there was a positive correlation between the emotion variables, which was significant (r=0.46, p < 0.001) (**H_1**). For the primary LIWC variable, affect, there was also a positive correlation between the articles and the tweets (r=0.21, p < 0.01) (**H_1**). Positive emotion in the articles also has a small positive correlation with negative emotion in the tweets (r=0.27, p < 0.001) (**H_1**). This is an interesting result here because as the tweets have a more negative emotion, this seems to be correlated with articles that have more positive emotion. These results reflect the research by Lee and Hamilton (2022) that found that there was more emotion in tweets than articles by journalists.

4.3 Correlation of Tone Variables

Looking at more variables, we can also analyse the tone variables as well as the positive tone and negative tone variables that are found under the affect section of the results in LIWC. There are positive correlations for Pearson's r with varying significant levels.

Table 4. Correlation Matrix for Tone Variables

Pearson's Correlations

Variable		**A - Tone**	**A - Positive Tone**	**A - Negative Tone**	**T - Tone**	**T - Positive Tone**	**T - Negative Tone**
1. A - Tone	Pearson's r	—					
2. A - Positive Tone	Pearson's r	0.70 ***	—				
3. A - Negative Tone	Pearson's r	-0.67 ***	-0.07	—			
4. T - Tone	Pearson's r	0.31 ***	-0.04	-0.41 ***	—		
5. T - Positive Tone	Pearson's r	0.36 ***	0.18 *	-0.29 ***	0.69 ***	—	
6. T - Negative Tone	Pearson's r	-0.17 *	0.09	0.27 ***	-0.42 ***	-0.24 ***	—

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

There is a moderate positive correlation found between the tweets and the articles for the overall tone (r=0.31, p < 0.001); see Table 4. Positive tone has a small positive correlation strength between tweets and articles (r=0.18, p<0.05). Negative tone also has a small positive correlation between tweets and articles (*r*=0.27, $p < 0.001$). Interesting results that can be seen in the table are that negative tone in tweets and tone in articles have a small negative correlation (*r*=-0.17, $p<0.05$). Meaning that the more negative tone tweets, the less overall tone in articles. And the higher the tone in articles, the lower the negative tone in tweets. This makes sense, however, when considering that tone is a summary variable, and the higher the tone, the more positive the tone is. This also describes the result found in the positive tone in tweets and the negative tone in articles with (*r*=-0.29, $p<0.001$). So as the tweets are more positive in tone, the articles are less negative in their tone.

A larger pairwise correlation with all variables of emotion and tone together (A -tone, A-positive tone, A - negative tone, A – affect, A-emotion, A-negative emotion, A-positive emotion, T-tone, T- positive tone, T- negative tone, T– affect, T-emotion, T-negative emotion, T-positive emotion) can be found in the Appendix in Table 5.

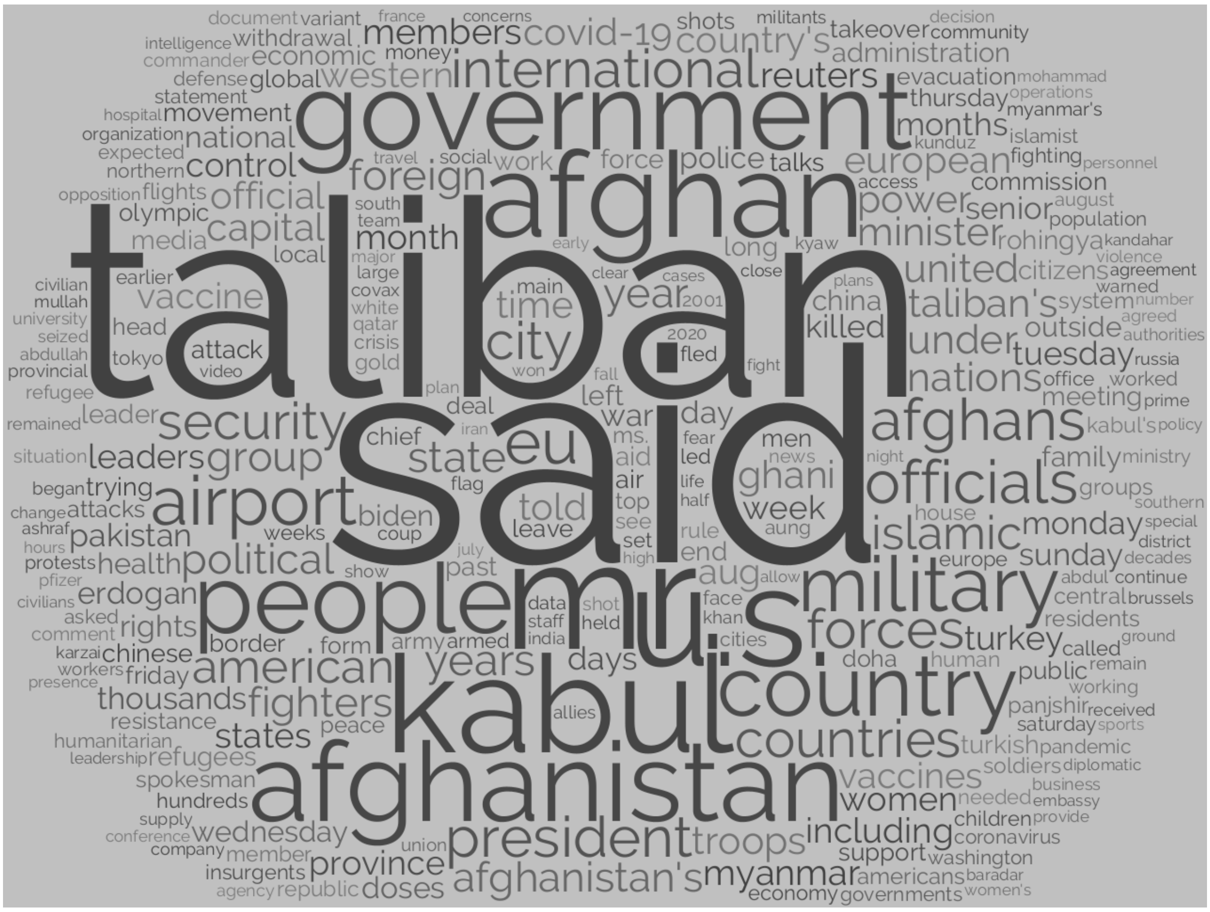
4.4 Correlation for Further Variables (Cognition, Social, Conflict, Moral, Culture and Political)

More variables of interest here are the correlation between A (articles) and T (tweets) for political, culture, conflict, cognition, social, and moral words. The correlation between articles and tweets for political, culture, conflict, and cognition words were positively correlated. These results can be seen from t=Table 6 in the Appendix of the study. The small positive correlation for political was Pearson's r (*r*=0.26, $p < 0.001$). The small positive correlation for culture was Pearson's r (*r*=0.18, $p <0.01$), while the moderate positive correlation for conflict was Pearson's r (*r*=0.36, $p < 0.001$). Lastly, cognition had a small positive correlation with Pearson's r (*r*=0.25, $p <0.001$). These results can be seen in Table 6 in the Appendix of the study. However, both moral words and social words in articles and in tweets were not correlated with each other, respectively. The table shows interesting results regarding the correlation between conflict words in articles and moral words in articles (*r*=0.27, $p<0.001$). Articles with conflict words also had a small positive correlation with tweets with political words (*r*=0.26, $p<0.001$).

Although there were some outliers in the dataset, we do not believe that there was any incorrect input but rather that this occurrence was a consequence of using LIWC in the first place. For our data, we also found that it was not normally distributed by doing Q plots but also saw this outcome as a result of the results garnered from LIWC. These Q plots can be seen in Appendices 1b, 1c and 1d.

4.5 Word Frequencies in Articles and Tweets

Figure 4. Word Cloud of Word Frequencies – Articles



The word cloud above shows the abundance of words that were in the articles that were analysed. The word “said” is at the top, was used 2481 times. But it is clear that since many of the articles were written by journalists focused on the withdrawal of the United States of America’s military from Afghanistan, the “Taliban” is at the top was written 2250 times, with “Taliban’s” also written 217 times. “Kabul” showed up 984 times while “Afghan” showed up 840 times, “Afghanistan” 771 times, and “government” 752 times.

Figure 5. Word Cloud of Word Frequencies –Tweets

The most-tweeted word in the dataset was “https”, tweeted 1381 times, showing that journalists often linked other articles. Without “https”, the most frequently tweeted word was also “Taliban”, tweeted 594 times, followed by “Kabul” and “Afghanistan” tweeted 389 times and 315 times, respectively.

5. Discussion

This study took inspiration from a wide variety of fields to try to provide insight into questions about correlating emotional content expressed by journalists and articles and in personal tweets. The findings of the study show that there is a positive correlation in the overall emotional content between the tweets and the articles written by the 6 journalists. This supports our hypothesis and adds to prior research from Lee and Hamilton (2022) that found higher usage of emotion words in tweets than in articles. This was supported by means of affect, overall emotion, negative emotion, positive emotion, negative tone, and positive tone for both tweets and articles. However, this was not the case for tone, as the mean for tone was larger in articles (18.48%) than for tweets (14.39%). This is an interesting result as it goes against prior research and maybe is due to the difference between the summary tone variable in LIWC and the emotion words variable. However, there were a large number of positive correlations found here that were significant. There were significant positive correlations between articles and tweets for the data points considered for negative emotion, emotion, affect, negative tone, positive tone, and tone. We can see that even though text analysis is one tool, there is a pattern here with text that may imply that there is a correlation of emotion between the two sources. It is important, however, not to draw too broad conclusions from the study. The study focuses on text analysis which looks only at the text of individuals and does not actively engage the individual.

The maximum article with overall emotion was "***Cold, Frightened and Armed: In Myanmar's Jungles, a Struggling Resistance***" by journalist Hannah Beech. Keywords that could be determined to have emotion were words like "hopes" and "empathize" which showed up 2 times and 1 time, respectively, in the article and are also in the LIWC dictionary. While for the article "***The U.S. edges China for the most gold medals***" also written by Hannah Beech there was the highest tone of 95.61%, while it had 0.2% emotion words only. An example of a concatenation of 3 tweets is: "Been there myself - although not three! Get well soon - and do lots of physio even if it hurts. https://t.co/lxU3K3lUti Just a reminder for those interested... https://t.co/UJOxnHTTyz https://t.co/mpSGXoLNzi". This concatenation had overall emotion words of 10.26% focused on the negative emotion words. The affect of this concatenation was 15.38%. Keywords in this tweet that holds emotion are "well" and "hurts"

This study also focused on political reporting and, in fact, explored mostly journalists that wrote in the context of conflicts. For some, they were focused on the crisis in Afghanistan, while others focused more on Myanmar in the time period. This is clear to see with the word frequencies in both the tweets and the articles being focused on the words "Taliban", "Kabul", "Afghanistan", "Afghan", and "government". This was the same for tweets which were focused on the words "Taliban", " Kabul", "Afghanistan", and "Afghan" were the top words. The focus on Afghanistan here is clear. A possibility with this result could also be that some journalists link their own articles or articles of fellow journalists in their tweets, which could lead to this correlation. This is also possibly seen with the 1380 times that "https" showed up in the concatenation of tweets.

Hence, it is also not so surprising to see the variables of conflict and political have positive correlations between the tweets and the articles, but this is not the case for moral words. This is interesting as in the research by Pantti (2019) she found an abundance of tweets focused on moral discourse in tweets. But this research does partly support articles and research presented by Pantti (2019) which explored how photo-journalists used Twitter as a place to share photos of the conflicts they were reporting on. Pantti (2019) found high use of subjectivity and reflected the moral claims of the journalists on Twitter, as well as high use of emotions and opinions that would not be so widely used in traditional journalism (Pantti, 2019). The correlations between moral words in tweets and articles, however, are not supported by this study though in contrast to Pantti's (2019) study.

The study is relevant because it is novel in its focus on the correlation of emotions in tweets and articles, which is the first that I know of in the field. Most journalists have Twitter as it is an easy way to get information out to their audience. Digital technological technologies like Facebook and

Twitter have allowed for the spreading of mass disinformation, which often also makes it into our journalism. The study asks what it means to be objective and if, in journalism, there is such a thing? While research from Lee and Hamilton (2022) found that there was a higher usage of system 1 thinking in tweets than in emotions, which was also found in this study. This article adds to the field by focusing on whether there is a correlation between the two. This study also shows the power of LIWC and the bags-of-words approach to find the amount of emotion within the text.

5.1. Limitations

This paper only made use of the LIWC text analysis programme, which has a bag-of-words approach and, as such, can miss some of the contexts in the text (Lee & Hamilton, 2022). As such, the programme can possibly miss the context of the larger text. Furthermore, as was mentioned in the introduction, research from Van Hee, Lefever, and Hoste (2017) found that text analysis programmes focus on the explicit meaning of the words. So, as the focus is not on the implicit meaning and the context was not wholly considered, there is a possibility that a larger meaning was missed here. Another limitation is the scope of the research. As the study is a Master's thesis, there will always be obstacles regarding the scope. Here, the scope of the research is limited to six journalists. Consequently, this study only examines one type of journalist, foreign affairs journalists. There may be different results when it comes to covering journalists from other topics such as sports journalists or more culture-focused journalists. The study also only focused on journalism in English, so any correlations in any other language are not considered, and as such, there may be different results when analysing that journalism.

A LIWC limitation can be seen from the tweet, "Why a veteran lawyer's losing record makes him proud. https://t.co/P0iInmDqbD" where the overall emotion words percentage is 14.29, while positive emotion and negative emotion are both 7.14 each. The tone, however, here is 1. This is an interesting result because the sentiment is very low and negative. However, the lawyer here is also made proud by his losing record. This is interesting because "proud" fits in the overall emotion category and positive emotion category. This is a limitation of LIWC as the sentiment variable, tone, overlooks the context of the sentence and the title of the article.

While the research only examined tweets from journalists, and no retweets, it is still possible that there may be a correlation due to journalists linking to their articles or journalists linking to other articles. The research showed that "https" was written 1381 times in the dataset showing that there is a lot of linking to other articles. The personal branding part of Twitter is referred to by Olausson (2017) can be a problem because the tweet may have topic words that directly correlate to the article. A hypothesis can be formed that there is a correlation here as the research found a correlation of Pearson's r (r=0.26, $p < 0.001$), for political words. However, while these are limitations, they can also be avenues for future work. While this research only analysed Western media outlets, this research could spur further research into work not focused on Western media outlets. It may be interesting and more striking to examine countries where there may be repression of press freedom as these journalists are constantly living and working under repressive conditions.

5.2. Behavioural Science Implications

Behavioural science was a crucial field that was researched for the present study. The study sprouted from interests in decision-making and objectivity and what it means to be objective. Peters et al. (2006) found that incidental affect has an impact on decision-making in people. This event can directly lead to us making decisions that are impacted by our emotions whether that is in professions such as journalism or a profession such as finance. As George and Dane (2016) found that emotions have an impact on people's investment choices as well as their bidding choices, it is clear that emotion must be given more insight in the decision-making field. Lee and Hamilton's (2022) study on system 1 thinking found that our tweets are more prone to emotion and, as such, can be prone to more cognitive bias. This is something to explore as they may relate to articles being written by the same journalists.

The implications of this research expand to Artificial Intelligence and the development of Artificial Intelligence tools to be used by humans. If we are to develop AI tools to work in collaboration with humans, we will need tools that are able to understand the human capacity to make decisions as a whole rather than focusing squarely on viewing humans as only rational and objective. It is imperative that AI tools be able to work better with humans in all forms as we charge towards a society that sees AI and humanity working next to each other to be productive.

5.3. Implications for Journalism

Journalism finds itself at a tricky junction (Baughman, 2015). As the decline of traditional journalism continues, journalists and media outlets are trying to find a place for them and their work to continue. While there is a lamenting by the mainstream media about the decline of traditional and their battle to protect the halls of serious journalism, there has been a democratisation of information. After World War II, as journalism started to become more critical of the government as they became more independent from government, they began to be snapped up by singular owners or by corporations. However, this did not stop the decline of traditional journalism with the blossoming of the internet. The internet has allowed for the rise of citizen journalists as well as data journalists to have access to a wider audience. Twitter allows for the ability of these smaller digital outlets to grow their audience. So, while we are presumed to be in the decline of journalism still, I believe that we are actually in a transition phase to a new journalistic method, one that may rely more on the emotional side as well.

6. Conclusion

This research found that there are correlations between tweets and articles by journalists. It builds on findings from Lee and Hamilton (2022), which found that journalists use more emotion words in their tweets than in their articles. This study is novel to my knowledge as this is the only research that is looking at the correlation between writing on Twitter and writing for articles. So even though there are caveats underlying this research, this study nonetheless provides a foundation for further research in this field. From this research, other researchers can expand this dataset to explore how tweets correlate with articles on a larger scope and for other forms of journalism such as sports journalism, culture journalism, or even tabloid journalism.

7. References

1. Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artificial Intelligence Review*, *54*(8), 5789–5829. https://doi.org/10.1007/s10462-021-09958-2
2. Ad Fontes Media. (2022). Interactive Media Bias Chart. Retrieved February 20, 2022, from https://adfontesmedia.com/interactive-media-bias-chart/
3. Baughman, J. (2015-09-01). The Decline of Journalism since 1945. In Making News: The Political Economy of Journalism in Britain and America from the Glorious Revolution to the Internet. : Oxford University Press. Retrieved 19 Jun. 2022, from https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199676187.001.0001/acprof-9780199676187-chapter-7
4. Beckett, C. (2015). How journalism is turning emotional and what that might mean for news. *POLIS: journalism and society at the LSE*.
5. Beckett, C., & Deuze, M. (2016). On the Role of Emotion in the Future of Journalism. *Social Media + Society*. https://doi.org/10.1177/2056305116662395
6. Bird, S. E. (2009). The future of journalism in the digital environment. *Journalism*, *10*(3), 293–295. https://doi.org/10.1177/1464884909102583
7. Bleiker, R., & Hutchinson, E. (2008). Fear no more: Emotions and world politics. *Review of International Studies, 34*(S1), 115-135. doi:10.1017/S0260210508007821
8. Bollen, J., Pepe, A., & Mao, H. (2009). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *arXiv preprint arXiv:0911.1583*.
9. Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proceedings of the international AAAI conference on web and social media* (Vol. 5, No. 1, pp. 450-453).
10. Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. https://doi.org/https://doi.org/10.1016/j.jocs.2010.12.007
11. Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). LIWC-22: Descriptive Statistics and Norms. https://www.liwc.app
12. Boyd, R., Ashokkumar, A., Seraj, S., & Pennebaker, J. (2022). *The Development and Psychometric Properties of LIWC-22*. https://doi.org/10.13140/RG.2.2.23890.43205
13. Canter, L. (2015). Personalised Tweeting. *Digital Journalism*, *3*(6), 888–907. https://doi.org/10.1080/21670811.2014.973148
14. Chatterjee, A., Narahari, K. N., Joshi, M., & Agrawal, P. (2019). SemEval-2019 Task 3: Emo Context Contextual Emotion Detection in Text. *Proceedings of the 13th International Workshop on Semantic Evaluation*, 39–48. https://doi.org/10.18653/v1/S19-2005
15. Christian, S. E. (2013). Cognitive Biases and Errors as Cause—and Journalistic Best Practices as Effect. *Journal of Mass Media Ethics*, *28*(3), 160–174. https://doi.org/10.1080/08900523.2013.794674
16. Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic Markers of Psychological Change Surrounding September 11, 2001. *Psychological Science*, *15*(10), 687–693. https://doi.org/10.1111/j.0956-7976.2004.00741.x
17. Dahlgren, P. (2018). Media, Knowledge and Trust: The Deepening Epistemic Crisis of Democracy. *Javnost - The Public*, *25*(1–2), 20–27. https://doi.org/10.1080/13183222.2018.1418819
18. De Choudhury, M., Counts, S., & Gamon, M. (2012). Not All Moods Are Created Equal! Exploring Human Emotional States in Social Media. *Proceedings of the International AAAI*

Conference on Web and Social Media, *6*(1), 66-73. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14279

19. De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting Depression via Social Media. *Proceedings of the International AAAI Conference on Web and Social Media*, *7*(1), 128-137. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14432
20. Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, *98*(45-60), 16.
21. Entman, R. M., & Gross, K. A. (2008). Race to Judgment: Stereotyping Media and Criminal Defendants. *Law and Contemporary Problems*, *71*(4), 93–133. http://www.jstor.org/stable/27654685
22. Feldman, R. (2013). Techniques and applications for sentiment analysis: The main applications and challenges of one of the hottest research areas in computer science. *Communications of the ACM*, *56*(4), 82–89. https://doi.org/10.1145/2436256.2436274
23. Fahmy, S., Taha, B. & Karademir, H. (2022). Journalistic Practices on Twitter: A Comparative Visual Study on the Personalization of Conflict Reporting on Social Media. *Online Media and Global Communication*, *1*(1), 23-59. https://doi.org/10.1515/omgc-2022-0008
24. George, J. M., & Dane, E. (2016). Affect, emotion, and decision making. *Organizational Behavior and Human Decision Processes*, *136*, 47–55. https://doi.org/https://doi.org/10.1016/j.obhdp.2016.06.004
25. Hasan, M., Rundensteiner, E., & Agu, E. (2019). Automatic emotion detection in text streams by analyzing Twitter data. *International Journal of Data Science and Analytics*, *7*(1), 35–51. https://doi.org/10.1007/s41060-018-0096-z
26. Hedman, U., & Djerf-Pierre, M. (2016). Journalists' uses of Twitter. In B. Franklin, & S. A. Aldridge II (Eds.), *The Routledge Companion to Digital Journalism Studies* (pp. 417-426). Routledge. https://doi.org/10.4324/9781315713793
27. Hermida, A. (2016). Twitter, Breaking The NEWS, and Hybridity in Journalism. In Franklin, B., & Eldridge, S.A. (Eds.). (2016). *The Routledge Companion to Digital Journalism Studies* (pp. 407-416) (1st ed.). Routledge. https://doi.org/10.4324/9781315713793
28. Holtzman, N. S., Tackman, A. M., Carey, A. L., Brucks, M. S., Küfner, 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
29. Howes, C., Purver, M., & McCabe, R. (2014). Linguistic Indicators of Severity and Progress in Online Text-based Therapy for Depression. *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 7–16. https://doi.org/10.3115/v1/W14-3202
30. Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, *8*(4), 407–424.
31. Kaur, J., & R. Saini, J. (2014). Emotion Detection and Sentiment Analysis in Text Corpus: A Differential Study with Informal and Formal Writing Styles. *International Journal of Computer Applications*, *101*, 1–9. https://doi.org/10.5120/17712-8078
32. Lancashire, I., & Hirst, G. (2009). Vocabulary Changes in Agatha Christie's Mysteries as an Indication of Dementia: A Case Study. *19Th Annual Rotman Research Institute Conference, Cognitive Aging: Research and Practice*.
33. Lasorsa, D. L., Lewis, S. C., & Holton, A. E. (2012). Normalizing Twitter. *Journalism Studies*, *13*(1), 19–36. https://doi.org/10.1080/1461670X.2011.571825
34. Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. *Literary and Linguistic Computing*, *26*(4), 435–461. https://doi.org/10.1093/llc/fqr013
35. Lee, J., & Hamilton, J. T. (2022). Anchoring in the past, tweeting from the present: Cognitive bias in journalists' word choices. *PLOS ONE*, *17*(3), e0263730. https://doi.org/10.1371/journal.pone.0263730

36. McDonnell, M., Owen, J. E., & Bantum, E. O. (2020). Identification of Emotional Expression With Cancer Survivors: Validation of Linguistic Inquiry and Word Count. JMIR formative research, 4(10), e18246. https://doi.org/10.2196/18246
37. Moeller, S. D. (2004). Media coverage of weapons of mass destruction.
38. Monzani, D., Vergani, L., Pizzoli, S. F. M., Marton, G., & Pravettoni, G. (2021). Emotional Tone, Analytical Thinking, and Somatosensory Processes of a Sample of Italian Tweets During the First Phases of the COVID-19 Pandemic: Observational Study. J Med Internet Res, 23(10), e29820. https://doi.org/10.2196/29820
39. Nielsen, R.K. (Ed.). (2015). *Local Journalism: The Decline of Newspapers and the Rise of Digital Media*. London • New York: I.B. Tauris. Retrieved June 20, 2022, from http://dx.doi.org/10.5040/9780755695171
40. Noguera, J. (2013). How open are journalists on Twitter? Trends towards the end-user journalism. *Communication & Society*, *26*, 93–114.
41. Olausson, U. (2017). The Reinvented Journalist. *Digital Journalism, 5*(1), 61–81. https://doi.org/10.1080/21670811.2016.1146082
42. Orgeret, K. S. (2020). Discussing Emotions in Digital Journalism. *Digital Journalism, 8* (2), 292–297. https://doi.org/10.1080/21670811.2020.1727347
43. Pantti, M. (2019). The Personalisation of Conflict Reporting. *Digital Journalism, 7* (1), 124–145. https://doi.org/10.1080/21670811.2017.1399807
44. Pantti, M. (2010). The value of emotion: An examination of television journalists' notions on emotionality. *European Journal of Communication*, *25*(2), 168–181. https://doi.org/10.1177/0267323110363653
45. Park, M., McDonald, D., & Cha, M. (2021). Perception Differences between the Depressed and Non-Depressed Users in Twitter. Proceedings of the International AAAI Conference on Web and Social Media, 7(1), 476-485. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14425
46. Peters, C. (2011). Emotion aside or emotional side? Crafting an 'experience of involvement' in the news. *Journalism*, *12*(3), 297–316. https://doi.org/10.1177/1464884910388224
47. Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision making: a "hot" topic. *Journal of Behavioral Decision Making*, *19* (2), 79–85. https://doi.org/https://doi.org/10.1002/bdm.528
48. Sacerdote, Bruce and Sehgal, Ranjan and Cook, Molly, Why is All Covid-19 News Bad News? (November 2020). NBER Working Paper No. w28110, Available at SSRN: https://ssrn.com/abstract=3735677
49. Sell, J., & Farreras, I. G. (2017). LIWC-ing at a Century of Introductory College Textbooks: Have the Sentiments Changed? Procedia Computer Science, 118, 108–112. https://doi.org/https://doi.org/10.1016/j.procs.2017.11.151
50. Seyeditabari, A., Tabari, N., & Zadrozny, W. (2018). Emotion Detection in Text: a Review. ArXiv, abs/1806.00674.
51. Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T.-S., & Zhu, W. (2017). Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, {IJCAI-17}*, 3838–3844. https://doi.org/10.24963/ijcai.2017/536
52. Shen, J. H., & Rudzicz, F. (2017). Detecting Anxiety through Reddit. Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology --- From Linguistic Signal to Clinical Reality, 58–65. doi:10.18653/v1/W17-3107
53. Skovsgaard, M. (2014). A tabloid mind? Professional values and organizational pressures as explanations of tabloid journalism. *Media, Culture & Society*, *36*(2), 200–218. https://doi.org/10.1177/0163443713515740
54. 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
55. Taboada, M. (2016). Sentiment Analysis: An Overview from Linguistics. *Annual Review of Linguistics*, *2* (1), 325–347. https://doi.org/10.1146/annurev-linguistics-011415-040518

56. Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. Proceedings of the International AAAI Conference on Web and Social Media, 4(1), 178-185. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14009
57. Van Hee, C., Lefever, E., & Hoste, V. (2018). We Usually Don{'}t Like Going to the Dentist: Using Common Sense to Detect Irony on {T}witter. *Computational Linguistics*, *44*(4), 793–832. https://doi.org/10.1162/coli_a_00337
58. Wahl-Jorgensen, K. (2020). An Emotional Turn in Journalism Studies? *Digital Journalism*, *8*(2), 175–194. https://doi.org/10.1080/21670811.2019.1697626
59. Waisbord, S. (2018). Truth is What Happens to News. *Journalism Studies*, *19*(13), 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881
60. Wang, W., Hernandez, I., Newman, D. A., He, J., & Bian, J. (2016). Twitter analysis: Studying US weekly trends in work stress and emotion. *Applied Psychology: An International Review*, *65*(2), 355–378. https://doi.org/10.1111/apps.12065

8. Appendices

Table 5. Correlation Matrix for Affect Variables

Pearson's Correlations				
			Pearson's r	
A - Tone	-	A - Affect	-0.16	*
A - Tone	-	A - Positive Tone	0.70	***
A - Tone	-	A - Negative Tone	-0.67	***
A - Tone	-	A - Overall Emotion	-0.15	*
A - Tone	-	A - Positive Emotion	0.19	**
A - Tone	-	A - Negative Emotion	-0.36	***
A - Tone	-	T - Tone	0.31	***
A - Tone	-	T - Affect	1.47e-3	
A - Tone	-	T - Positive Tone	0.36	***
A - Tone	-	T - Negative Tone	-0.17	*
A - Tone	-	T - Overall Emotion	-0.09	
A - Tone	-	T - Positive Emotion	-0.02	
A - Tone	-	T - Negative Emotion	-0.09	
A - Affect	-	A - Positive Tone	0.53	***
A - Affect	-	A - Negative Tone	0.81	***
A - Affect	-	A - Overall Emotion	0.53	***
A - Affect	-	A - Positive Emotion	0.32	***
A - Affect	-	A - Negative Emotion	0.50	***
A - Affect	-	T - Tone	-0.39	***
A - Affect	-	T - Affect	0.21	**
A - Affect	-	T - Positive Tone	-0.16	*
A - Affect	-	T - Negative Tone	0.30	***
A - Affect	-	T - Overall Emotion	0.41	***
A - Affect	-	T - Positive Emotion	0.05	
A - Affect	-	T - Negative Emotion	0.39	***
A - Positive Tone	-	A - Negative Tone	-0.07	
A - Positive Tone	-	A - Overall Emotion	0.27	***
A - Positive Tone	-	A - Positive Emotion	0.44	***
A - Positive Tone	-	A - Negative Emotion	0.05	
A - Positive Tone	-	T - Tone	-0.04	
A - Positive Tone	-	T - Affect	0.18	*
A - Positive Tone	-	T - Positive Tone	0.18	*
A - Positive Tone	-	T - Negative Tone	0.09	
A - Positive Tone	-	T - Overall Emotion	0.27	***
A - Positive Tone	-	T - Positive Emotion	0.05	
A - Positive Tone	-	T - Negative Emotion	0.26	***
A - Negative Tone	-	A - Overall Emotion	0.42	***
A - Negative Tone	-	A - Positive Emotion	0.07	
A - Negative Tone	-	A - Negative Emotion	0.54	***

Pearson's Correlations			
			Pearson's r
A - Negative Tone	-	T - Tone	-0.41 ***
A - Negative Tone	-	T - Affect	0.12
A - Negative Tone	-	T - Positive Tone	-0.29 ***
A - Negative Tone	-	T - Negative Tone	0.27 ***
A - Negative Tone	-	T - Overall Emotion	0.28 ***
A - Negative Tone	-	T - Positive Emotion	0.04
A - Negative Tone	-	T - Negative Emotion	0.27 ***
A - Overall Emotion	-	A - Positive Emotion	0.69 ***
A - Overall Emotion	-	A - Negative Emotion	0.82 ***
A - Overall Emotion	-	T - Tone	-0.32 ***
A - Overall Emotion	-	T - Affect	0.26 ***
A - Overall Emotion	-	T - Positive Tone	-0.13
A - Overall Emotion	-	T - Negative Tone	0.32 ***
A - Overall Emotion	-	T - Overall Emotion	0.46 ***
A - Overall Emotion	-	T - Positive Emotion	-0.07
A - Overall Emotion	-	T - Negative Emotion	0.49 ***
A - Positive Emotion	-	A - Negative Emotion	0.20 **
A - Positive Emotion	-	T - Tone	-0.13
A - Positive Emotion	-	T - Affect	0.18 *
A - Positive Emotion	-	T - Positive Tone	0.07
A - Positive Emotion	-	T - Negative Tone	0.15 *
A - Positive Emotion	-	T - Overall Emotion	0.29 ***
A - Positive Emotion	-	T - Positive Emotion	0.07
A - Positive Emotion	-	T - Negative Emotion	0.27 ***
A - Negative Emotion	-	T - Tone	-0.33 ***
A - Negative Emotion	-	T - Affect	0.18 *
A - Negative Emotion	-	T - Positive Tone	-0.22 **
A - Negative Emotion	-	T - Negative Tone	0.30 ***
A - Negative Emotion	-	T - Overall Emotion	0.38 ***
A - Negative Emotion	-	T - Positive Emotion	-0.14
A - Negative Emotion	-	T - Negative Emotion	0.43 ***
T - Tone	-	T - Affect	-0.08
T - Tone	-	T - Positive Tone	0.69 ***
T - Tone	-	T - Negative Tone	-0.42 ***
T - Tone	-	T - Overall Emotion	-0.46 ***
T - Tone	-	T - Positive Emotion	-0.02
T - Tone	-	T - Negative Emotion	-0.45 ***
T - Affect	-	T - Positive Tone	0.26 ***
T - Affect	-	T - Negative Tone	0.88 ***
T - Affect	-	T - Overall Emotion	0.71 ***
T - Affect	-	T - Positive Emotion	0.22 **
T - Affect	-	T - Negative Emotion	0.66 ***
T - Positive Tone	-	T - Negative Tone	-0.24 ***
T - Positive Tone	-	T - Overall Emotion	-0.20 **
T - Positive Tone	-	T - Positive Emotion	0.31 ***
T - Positive Tone	-	T - Negative Emotion	-0.28 ***
T - Negative Tone	-	T - Overall Emotion	0.81 ***
T - Negative Tone	-	T - Positive Emotion	0.07
T - Negative Tone	-	T - Negative Emotion	0.81 ***
T - Overall Emotion	-	T - Positive Emotion	0.18 *
T - Overall Emotion	-	T - Negative Emotion	0.96 ***
T - Positive Emotion	-	T - Negative Emotion	-0.10
* p < .05, ** p < .01, *** p < .001			

Table 6. Correlation Matrix for Other Variables (Cognition, Conflict, Moral, Culture, Political)

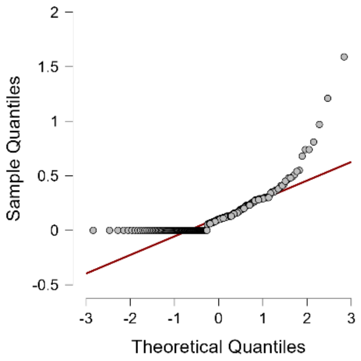
Pearson's Correlations		
		Pearson's r
A - Cognition	- A - Social	-0.19 **
A - Cognition	- A - Conflict	-0.33 ***
A - Cognition	- A - moral	-0.07
A - Cognition	- A - Culture	-0.16 *
A - Cognition	- A - Political	-0.13
A - Cognition	- T - Cognition	0.25 ***
A - Cognition	- T - social	-0.02
A - Cognition	- T - Conflict	-0.25 ***
A - Cognition	- T - Moral	-0.11
A - Cognition	- T - Culture	-0.16 *
A - Cognition	- T - political	-0.04
A - Social	- A - Conflict	0.52 ***
A - Social	- A - moral	0.47 ***
A - Social	- A - Culture	0.42 ***
A - Social	- A - Political	0.43 ***
A - Social	- T - Cognition	-0.29 ***
A - Social	- T - social	0.15 *
A - Social	- T - Conflict	0.12
A - Social	- T - Moral	-0.13
A - Social	- T - Culture	0.30 ***
A - Social	- T - political	0.09
A - Conflict	- A - moral	0.27 ***
A - Conflict	- A - Culture	0.37 ***
A - Conflict	- A - Political	0.35 ***
A - Conflict	- T - Cognition	-0.17 *
A - Conflict	- T - social	0.22 **
A - Conflict	- T - Conflict	0.36 ***
A - Conflict	- T - Moral	0.03
A - Conflict	- T - Culture	0.23 **
A - Conflict	- T - political	0.26 ***
A - moral	- A - Culture	0.24 ***
A - moral	- A - Political	0.27 ***
A - moral	- T - Cognition	-0.09
A - moral	- T - social	0.13
A - moral	- T - Conflict	0.09
A - moral	- T - Moral	0.06
A - moral	- T - Culture	0.07
A - moral	- T - political	-0.03
A - Culture	- A - Political	0.87 ***
A - Culture	- T - Cognition	-0.21 **
A - Culture	- T - social	0.09
A - Culture	- T - Conflict	0.15 *
A - Culture	- T - Moral	-0.10
A - Culture	- T - Culture	0.18 **
A - Culture	- T - political	0.17 *
A - Political	- T - Cognition	-0.19 **
A - Political	- T - social	0.11
A - Political	- T - Conflict	0.03
A - Political	- T - Moral	-0.09
A - Political	- T - Culture	0.21 **
A - Political	- T - political	0.26 ***
T - Cognition	- T - social	0.06
T - Cognition	- T - Conflict	-0.04
T - Cognition	- T - Moral	0.16 *
T - Cognition	- T - Culture	-0.62 ***
T - Cognition	- T - political	-0.21 **
T - social	- T - Conflict	0.41 ***
T - social	- T - Moral	0.16 *
T - social	- T - Culture	-0.11
T - social	- T - political	0.43 ***

Pearson's Correlations			
			Pearson's r
T - Conflict	-	T - Moral	0.16 *
T - Conflict	-	T - Culture	-0.04
T - Conflict	-	T - political	0.14
T - Moral	-	T - Culture	-0.13
T - Moral	-	T - political	-0.04
T - Culture	-	T - political	0.26 ***

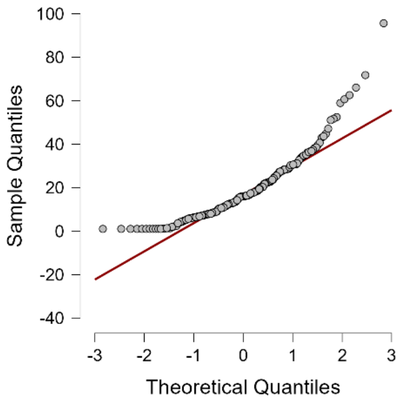
* p < .05, ** p < .01, *** p < .001

Appendix 1b. Q-Plots for emotion variables for articles

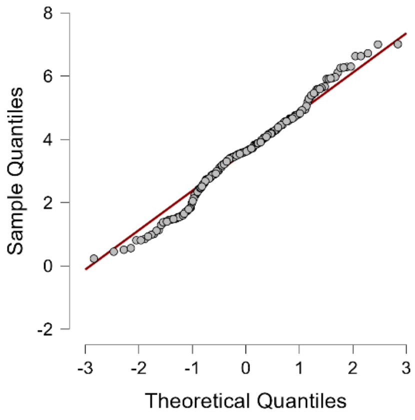
A - Positive Emotion



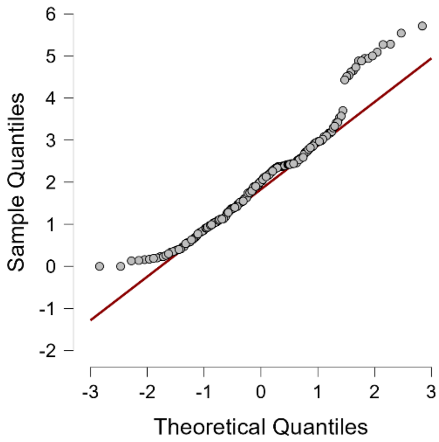
A - Tone



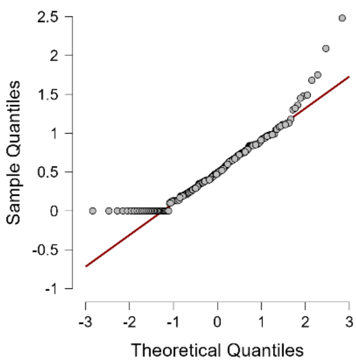
A - Affect



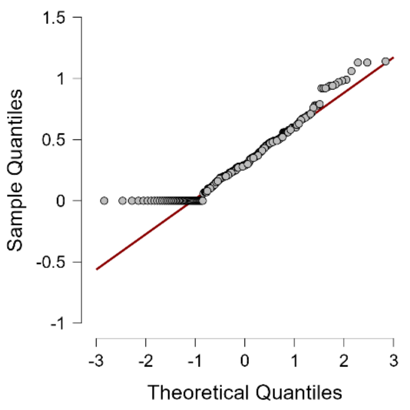
A - Negative Tone



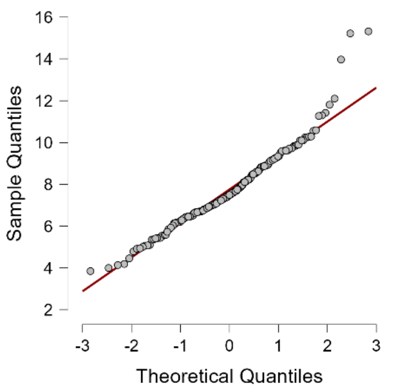
A - Overall Emotion



A - Negative Emotion

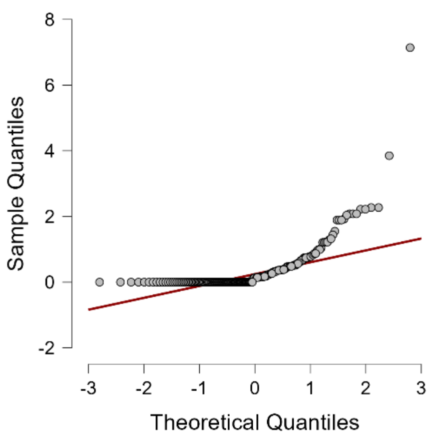


A - Cognition

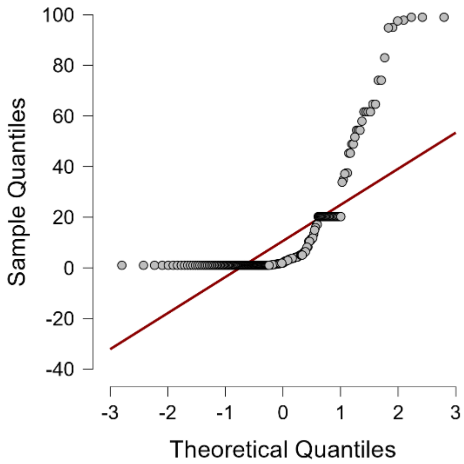


Appendix 1c. Q-plots for emotion variables in Tweets

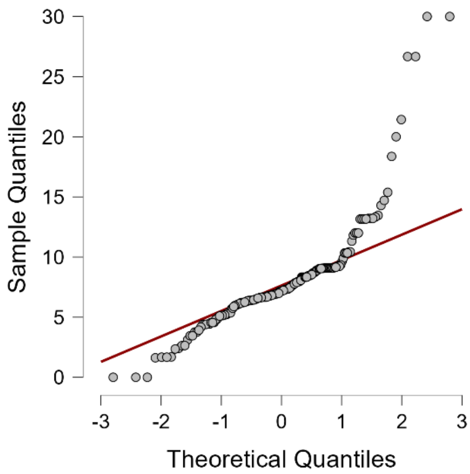
T - Positive Emotion



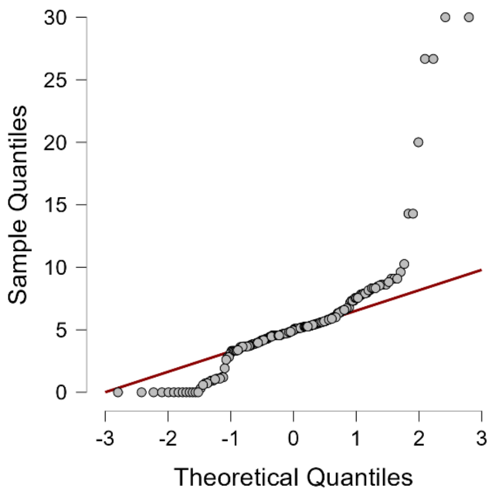
T - Tone



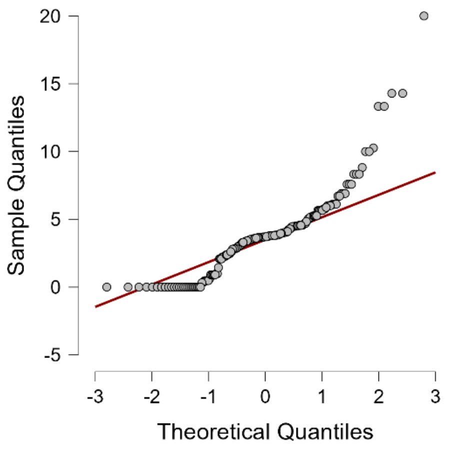
T - Affect



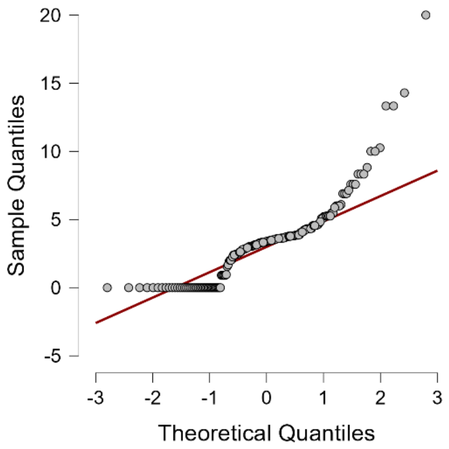
T - Negative Tone



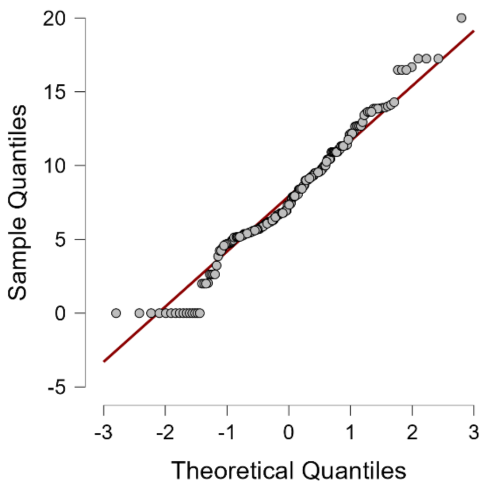
T - Overall Emotion



T - Negative Emotion

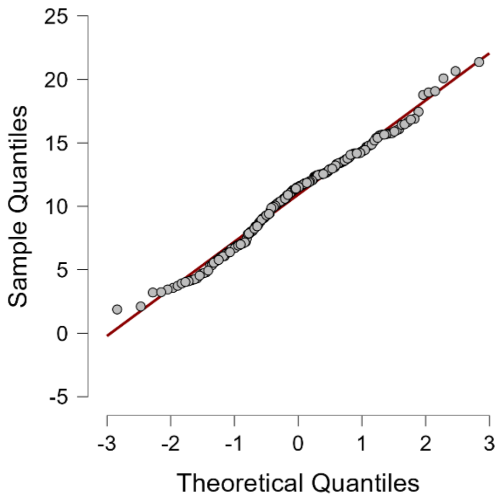


T - Cognition

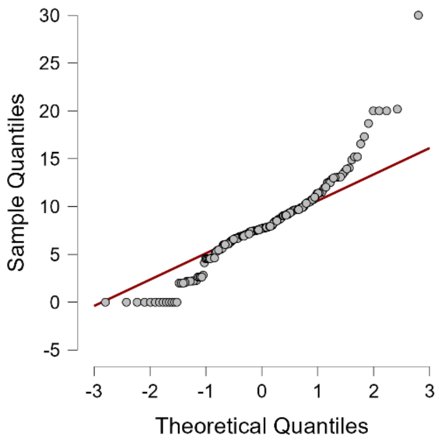


Appendix 1d – Q plots – Social, Conflict, Moral, Culture, Political variables

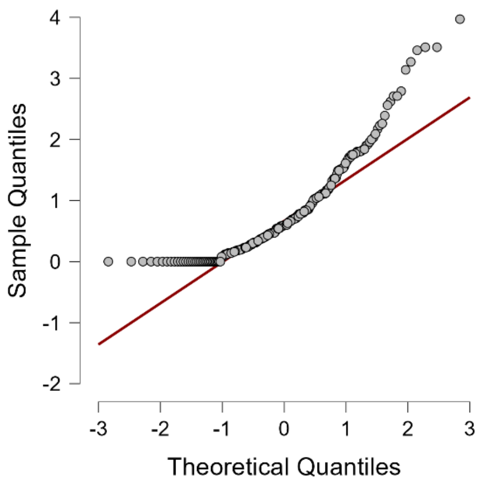
A - Social



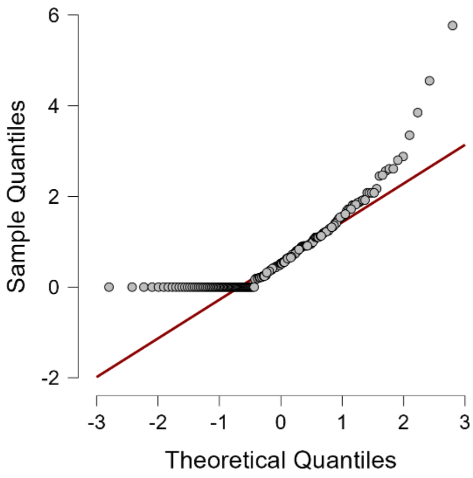
T - social



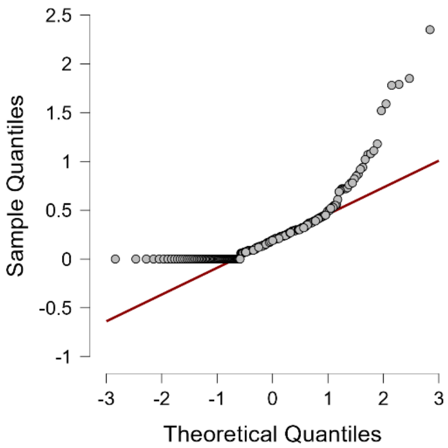
A - Conflict



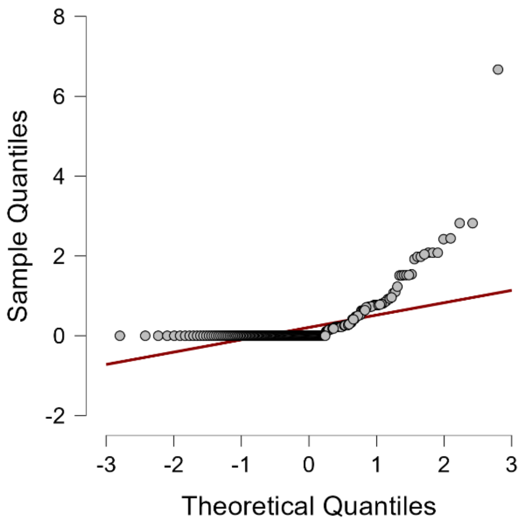
T - Conflict



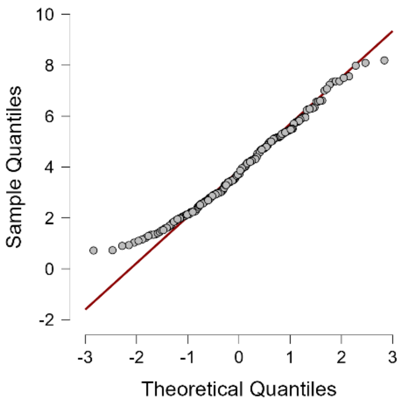
A - moral



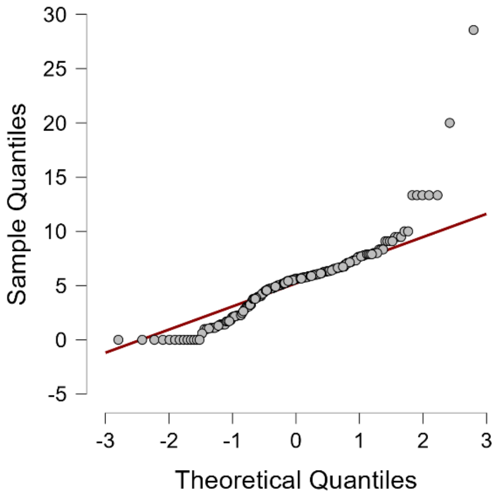
T - Moral



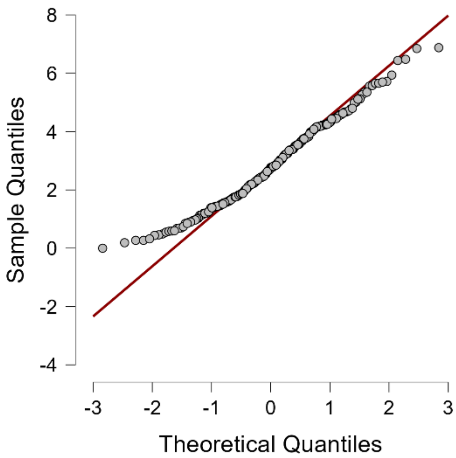
A - Culture



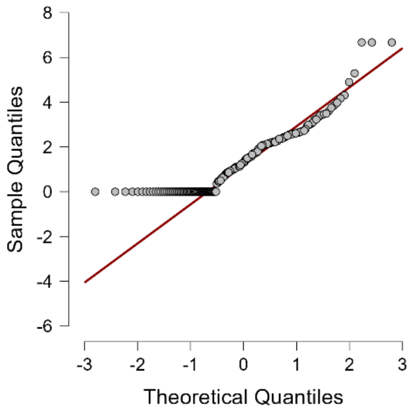
T - Culture



A - Political



T - political


Sample Quantiles
Theoretical Quantiles
8
6
4
2
0
-2
-4
-6
-3
-2
-1
0
1
2
3