



Person or PC? A Comparison of Human and Computer Coding as Content Analyses Tools Evaluating Severe Weather

Cory L. Armstrong^{1*}

0000-0003-4730-9665

Nathan A. Towery²

0000-0001-5511-0547

¹ University of Alabama, USA

² Jackson State University, USA

* Corresponding author: cory.l.armstrong@ua.edu

Citation: Armstrong, C. L., & Towery, N. A. (2022). Person or PC? A Comparison of Human and Computer Coding as Content Analyses Tools Evaluating Severe Weather. *Online Journal of Communication and Media Technologies*, 12(2), e202211. <https://doi.org/10.30935/ojgmt/11572>

ARTICLE INFO

Received: 30 Nov 2021

Accepted: 12 Jan 2022

ABSTRACT

Computer-aided content analyses programs have been deployed for social science research in recent years; however, few studies have evaluated their effectiveness, compared to human coding. This study uses open-ended responses from respondents seeking information in preparation for Hurricane Michael to compare human- and computer-coding. In particular, the comparison involves the use of Excel as a common and relatively simple coding instrument. Results indicated significant differences between frequencies coded by humans and a computer, with additional findings suggesting that residents employ television as a tool for information gathering when severe weather is imminent. Final discussion focuses on support for a blended model of both human and computer coding, while examining the findings related to severe weather.

Keywords: content analysis, methodology, computer-aided coding, disaster preparedness, information gathering

INTRODUCTION

The past few years have seen exponential growth in the use of computer-aided textual and content analyses programs, mainly to aid in humanities and social science scholarship. Quick online searches promote proprietary software programs including MaxQDA, Atlas.ti, Diction, and NVivo. Others promote open-source programs (e.g., Lexicoder) or even a basic frequency count via Excel. Conversely, most social science researchers are taught that to dig deep into content for meaning and interpretation, the use of human coding remains the gold standard, because computers are unable to determine nuances and underlying meaning within content. Of course, humans are fallible. Therein is the quandary. Which is better for media content studies—humans or computers?

The purpose of this study is to attempt to answer that question by comparing the use of individual or human coding on manifest content to examine the similarities and differences within that analysis. To investigate this topic, an open-ended dataset on how individuals gather information about imminent severe weather was employed, as the unpredictability and uncertainty of severe weather provides an important backdrop for information-gathering studies. The content under study provides a unique opportunity to see whether significant differences within human and computer coding exist in fairly straightforward, manifest content.

Further, the study will examine how individuals in the southeastern U.S. and in the path of Hurricane Michael described their information-gathering process about the storm in the days surrounding that time. The data for the study were collected immediately following the October 2018 land strike of Hurricane Michael. Thus, individuals had recently experienced a major hurricane, allowing for their responses to accurately describe what media and information sources they used to prepare for the storm. This project can provide unique insight into the way individuals describe their information gathering and, thus, decision-making process. That is, where do they go for information when they need credible information quickly? In addition, it will provide a clear comparison between manifest computer coding and nuanced human coding—evidence of which is sorely needed in a time of big data and media credibility.

LITERATURE REVIEW

One of the core social science methodologies for mass communication is content analysis. It is defined as such:

"the systematic and replicable examination of symbols of communications, which have been assigned numeric values according to valid measurement rules, and the analysis of relationships involving those values using statistical methods, to describe the communication, draw inferences about its meaning, or infer from the communication to its context, both of production and consumption (Riffe et al., 2014, p. 19).

Perhaps the three key elements of that definition are symbols, measurement, and meaning. Symbols refer to the text, image, video or other element within communication that will be studied (Riffe et al., 2014). Measurement refers to the way that each element is converted to a numeric value or category as part of the coding process (Riffe et al., 2014). Finally, and perhaps most importantly, is the process of using the symbols and the measurement to determine the meaning of the content, both on the surface (manifest) or under the surface (latent). Krippendorff (2004) discusses the various ways that meaning can be deduced, based on the context, societal norms and coders' experiences. In addition, there can be dominant and oppositional readings, suggesting that a main reading and a subtler reading can also be found.

Given the complexity in coding then, it follows that much content analysis has been conducted through human coders over the years. However, over the past 10-15 years, several computer-aided content analysis programs have been developed and employed in social science research, as noted earlier in this paper. It begs the question of how effective are computers at coding content?

Su et al. (2016) suggests that comparisons between computer and human coding falls into three facets: reliability, validity, and efficiency. Reliability focuses on the likelihood that a group of researchers would come to the same general conclusions on interpreting pieces of data as any individual scholar, using similar parameters (see Krippendorff, 2004). While variation could occur within humans by individual choice, which brings up the importance of statistical reliability testing, computer coding will have no such reliability issues, as the content count does not change.

Validity in content analysis is achieved when inferences and conclusions drawn on a piece of text are appropriate and have similarly-deduced meanings, even after new evidence and new observations are made (Krippendorff, 2004). This brings up the issue of manifest content, which is the basic frequency of occurrences of a symbol within a text, and latent content, which focuses on judgements and interpretations of underlying meanings within text. Clearly, in manifest content, computer coding should excel, whereas, human coding would prevail in examining latent content, as computers cannot denote meaning from language.

Finally, Su et al. (2016) argue that for large data sets, computer coding has a higher level of efficiency for coding. Computers can mine data in minutes for words, phrases and frequencies what would take humans hours, days and weeks to review individually. They proposed a hybrid model of analyses, where individual coders examine and classify a segment of large data that is then programmed into the computer program through an algorithm (see also Guo et al., 2016). Similarly, Krippendorff (2004) discusses a collaborative approach to content analyses, where humans translate written documents into categories and then set up

algorithms to code large data portions. All of these authors (and others, such as Lewis et al., 2013) note that there are merits to computer-aided coding.

Finally, we argue that in addition to examining the merits of computer versus human coding, the comparison also needs to have a reason for evaluation. That is, why does it matter? Certainly, accuracy is a crucial point of determination, as with all methodology. Guo et al. (2016) found significant differences between the two. But, aside from the basic tenets of reliability, validity and efficiency, scholarly research can and should have impact. In the present work, the results can be used to understand where individuals go to obtain information to make life-saving decisions. We argue that when it comes to making decisions about how best to prepare for severe weather, the results of the comparison are crucial, as we'll discuss below.

Decision-Making and Preparatory Actions During Severe Weather

Scholarship within severe weather focuses on disaster preparedness, or how informed and ready individuals feel to make decisions regarding their safety. That is, what kind of information do individuals seek out to make the decision to evacuate an area when severe weather is imminent? Most research has focused at how governmental or "official" information is distributed, along with word-of mouth advice (e.g., Landwehr et al., 2016; Linardi, 2016), while others looked more at how broadcast media information may be gathered and the use of visual elements, such as weather graphics, within those messages (see e.g., Armstrong et al., 2020; Armstrong & Towery, 2021). Studies have focused on weather and technological tools to examine potential hurricane routes and wind-speed changes (Daniels & Loggins, 2007; Landsea & Franklin, 2013). Sociological-based studies have examined graphics to document damage and casualties, along with intensity (Ash et al., 2014; Moor & Dixon, 2011; Wu et al., 2014; Zahran et al., 2013).

However, some scholars have argued that these technologies, while innovative, are often incomprehensible to general audiences. For example, Sattler et al. (2000) suggest that the information contained may not help residents make decisions and allow for significant misinterpretations by consumers. This misunderstanding can expose residents to increased danger and harm. Other researchers argue that media graphics need to be more accurate and provide specific information about what kinds of preparatory actions need to be done (Mileti & O'Brien, 1992; Mileti & Sorensen, 1987). A recent experiment by Armstrong and Towery (2021) compared different versions of a broadcast media message alerting residents to a hypothetical hurricane to determine which had the most impact—live shot footage, computer cone-of-uncertainty, or text-only. For those who had just been through a hurricane, live shots were the most effective to motivate respondents to act, while for residents who had not just been through a hurricane, they preferred the computer model.

None of the studies above, however, have tracked the type of information that residents seek out when severe weather is imminent. Further, the mixed messages obtained from earlier studies do not provide conclusive information about what kind of information residents need to prepare for severe weather. The current study should shed light on the information sources that individuals seek out—information that may or may not persuade them to take preparatory action.

Preparatory actions

Most scholarship studying disaster preparedness focuses on one's likelihood to evacuate the area when severe weather (e.g., hurricanes) strikes. Wachinger et al. (2013) examined factors that may influence one to leave the area, finding that the level of trust individuals have in authority (such as police and fire) may influence their decision to evacuate. Others believe in an optimism bias, that negative consequences are unlikely to impact them in bad weather (Sheldon & Antony, 2018).

While scholars often focused on likelihood to evacuate the area, Kang et al. (2007) and, more recently, Armstrong et al. (2020) suggest that a broader series of preparatory actions are more accurate in understanding how individuals prepare for a disaster: secure their home and leave; protect property; pack travel items; gather household members; prepare to leave work; and travel work to home (Kang et al., 2007, p. 894). Although individuals may not choose to evacuate during severe weather, they may engage in preparatory actions. As Baker (1991) suggests, there are some general assumptions that prior exposure and trust in authority would be influential in predicting evacuation, it is difficult to discern their level of effectiveness in actual events.

To address this issue, the current study asks residents to describe the sources in which they sought out to prepare for Hurricane Michael and examines the impact those sources had on their likelihood to engage in preparatory actions. These descriptions will be coded in two ways to better understand the relationship between media and behavioral intention—and perhaps provide a level of accuracy in their descriptions that prior studies have not been able to capture. Therefore, based on the above literature, the following research questions are proposed.

RQ1: What are the relationships between information-gathering variables that were individually coded on one's likelihood to engage in preparatory actions?

RQ2: What are the relationships between information-gathering variables that were computer coded on one's likelihood to engage in preparatory actions?

RQ3: What is the comparison between information-gathering variables that were individually coded with those that were coded via computer?

METHODS

To address these questions, we conducted a content analysis of data collected through answers within a quantitative survey. As part of the survey, respondents were asked an open-ended question describing "How did you gather information about Hurricane Michael. For example, did you watch television, search on the internet, talk to friends? How did you make a decision about what preparation actions (if any) you undertook as part of the storm?" The answers were coded using two forms of content analyses—human/individual coding and computer coding. Each procedure will be outlined below. These analyses focus on the content provided within those answers to evaluate the differences in the two coding schemes.

Survey Participants

Participants were recruited for this institutional review board-approved survey via a Qualtrics panel (i.e., a pool of U.S. adults who have volunteered to participate in online survey research via the company). From November 30 to December 15, 2018, the study was in the field, and responses were anonymous and confidential. No personal identifiers were linked to participants.

Study participants ($N=567$) were drawn from residents within and near the path of Hurricane Michael: Mobile County (2016 population: 413,000) and Baldwin County (2016 population: 212,000) in Alabama; and in Florida, Escambia County, (2016 population: 313,000), Bay County (180,000), Santa Rosa County (166,000), Walton County (63,000), Leon County (285,000), Okaloosa County (197,000), and Wakulla County (31,586). The sample had 68 percent female participants, a median age of 55-64 years of age, and 83.6 percent chose white as their race. Roughly 50 percent of participants reported that they have lived in the same area for more than nine years and 90 percent of participants reported that they had prior experience with severe weather or a natural disaster. Specifically, 46.2 percent of participants reported suffering damage to their home or property from Hurricane Michael.

Dependent variable

The variable preparatory actions were created from a seven-item summative index asking respondents on a 1-7 scale which actions they would be likely to conduct if an actual hurricane was imminent: prepare to leave the area, evacuate current location and go home, gather household members, pack supplies for travel, protect property, secure home and evacuate the area, or seek shelter immediately ($M=38.58$, $SD=8.15$; Chronbach's alpha=.78). This scale is adapted in part from that of the (Morss et al., 2016) examination of preparatory actions.

Content Analysis Methods

Individual/human coding

Coding was conducted by two graduate students, with training by the first author. After initial training, coders were given 20 randomly chosen responses from the sample for initial coding. Upon discussion and resolution of disagreements from those stories, an additional 36 randomly selected responses were added,

along with 10 purposely chosen stories were added to this mix to ensure that coding for all key variables would be tested during the reliability coding phase. The authors discussed any final disagreements and came to a consensus on those issues. Final reliability was determined employing all 66 stories (11.6% of the final sample). Individual reliability scores will be discussed below.

Information sources were coded for the presence of each within the answers: television (82%, Krippendorff's alpha=.82), radio (17%, $\alpha=1.0$), internet, but not social media (40.1%, $\alpha=.93$), social media (14%, $\alpha=1.0$), friends/family (30.7%, $\alpha=1.0$), apps (14.8%, $\alpha=.89$), or newspaper (1.3%, $\alpha=1.0$). Coders were asked to list which source was mentioned first in their answer ($\alpha=.92$): television (72.1%); internet (7.3%) and friends/family (6.1%). *Evacuation* was coded as any mention of evacuation for Hurricane Michael (6.3%) or whether they evacuated (5.9%), compared to no mention at all (87%, $\alpha=.88$ for the entire variable). Finally, coders were asked to determine the main focus of the response: focus on evacuation 8.4%, $\alpha=.81$), focus on information gathering (94.8%, $\alpha=1.0$), focus on preparation (19.1%, $\alpha=.72$), or focus on safety (1.8%, $\alpha=1.0$),

Computer coding

This study also utilized a computer coding content analysis. The same sample of 567 responses were used in the analysis. All of the computer coding was done in Microsoft Excel and utilized different functions and key words. Excel was chosen to provide a simple frequency count of words. While other software programs offer more powerful analyses, including frame and theme analysis from content (e.g., MaxQDA or NVivo), this analysis was designed to provide a clear frequency count as compared to human coders. As noted by Bowmans and Trilling (2016), Excel is appropriate for basic frequency counts and visibility analysis.

Procedures: The responses used in the computer coding were assigned a unique identification number so that each response could be uniquely identified within the analysis. To identify manifest codes within the responses, the search function was used (e.g., =SEARCH("facebook", B:B)). The search function identified if that word was present, but also when it appeared in the response. For example, if "facebook" appeared in the response, it might indicate that Facebook appeared at the 67th character. This allowed the researchers to identify in which order certain words appeared within the responses. In order to identify which codes appeared first, second, and third, the small function was used (e.g., =SMALL(D3:S3, 1)). This function will identify the smallest number in a given array. In this case, the smallest number would indicate the first appearance of a code.

Lastly, the index and match functions were used to identify the correlating character number to the corresponding variable header (e.g., =INDEX(label, MATCH(T3, D3:S3,0))). Prior to this step, all we had were the three smallest numbers in a given response. This step identified which code was associated with which number. For example, if a response generated a code of 78 for television, and the small function identified 78 as the second smallest number, this function would identify that the 78 is associated with the variable header "television." These three functions were the only functions used in Microsoft excel for the computer coding procedures.

Description of variables: The computer coding analysis used many of the same variables that the human coding content analysis portion of the study utilized. Due to the latent nature of some of the variables in human coding content analysis, they were excluded from the computer coding. The search criteria for the computer coding was determined after the human coding content analysis was completed, as to understand the common manifest words respondents used for each variable. For all of the variables, only one of the search criteria had to be met in order for it to be counted as present.

The same seven information sources were used in the computer content analysis. For the presence or absence of television (presence=50.7%) as an information source, the key terms used were "television" and "tv." For the presence or absence of radio (presence=16.3%) as an information source, the key term used was "radio." For the presence or absence of the internet (presence=36.8%) as an information source, the key terms used were "internet," "website," ".com," and "online." For the presence or absence of social media (presence=13.2%) as an information source, the key terms used were "social media", "facebook", and "twitter." Other specific platforms were omitted due to the lack of presence from them in the human content analysis. For the presence or absence of friends or family (presence=27.9%) as an information source, the key terms used were "friend", "family", and "worker." For the presence or absence of a phone app (presence=19.8%) as

an information source, the key terms used were "phone" and "app". Lastly, for the presence or absence of the newspaper (presence=2.0%) as an information source, the key terms used was "paper."

The order in which the respondents identified their information sources was determined as well. The first, second, and third information sources mention were recorded in that order, in order to determine where individuals were getting their information from first. The variable that identified if the participant was focused on evacuation used search criteria as well. For the presence or absence of evacuation (presence=9.7%), the key terms used were "evacuat" and "left." This search criterion allowed for all variations on the spelling and contexts of evacuation to be captured. The variable that identified if the participant was focused on preparations used search criteria as well. For the presence or absence of preparations (presence=11.6%), the key term used was "prepar." This search criterion allowed for all variations on the spelling and contexts of preparation to be captured.

The variable that identified if the participant was focused on safety used search criteria as well. For the presence or absence of safety comp (presence=1.6%), the key term used was "safe." This search criterion allowed for all variations on the spelling and contexts of safety to be captured. The variable that identified if the participant was focused on information gathering used search criteria as well. For the presence or absence of information gathering (presence=29.6%), the key terms used were "info", "document", "telephone", "talk", "conversat", and "cell." This search criterion allowed for many variations on the spelling and contexts of information gathering to be captured.

The analysis for this paper was conducted using SPSS 24. Before analysis was conducted, all three datasets (survey, human coding and computer coding) were merged into one SPSS dataset. This allowed for an evaluation of normality among the measurement, along with a descriptive check of all variables to ensure the measurements fit all underlying assumptions.

RESULTS

To examine **RQ1**, which examined the relationship of individually coded media variables on preparatory actions, independent-sample t-tests were conducted on the presence and absence of our seven media variables on the likelihood to engage in preparatory actions. As noted in **Table 1**, only the appearance of television ($M=38.97$ when present, versus $M=36.72$ when absent) in the answers were statistically significantly ($p<.05$) likely to increase the likelihood of preparatory actions. Conversely, the presence of newspapers had a negative relationship ($p<.05$) on preparatory actions ($M=31.71$ when present, versus $M=38.65$ when absent). No significant relationships were found when we examined whether television and internet were mentioned first in the answers, and given the low percentage of appearance of any other media, we did not conduct significance testing on those.

Independent sample t-tests were conducted on our individually coded variables of focus information gathering and focus on evacuation. Statistically significant findings indicated that when responses were focused on evacuation ($M=41.81$ when present, versus $M=38.27$ when absent) and on information ($M=38.80$ when present, versus $M=34.14$ when absent), the level of preparatory actions were higher than when they were absent from the response. Based on the individual coding results, it appears that television as an information source was a positive influence on preparatory actions, whereas newspapers were a negative influence. Further, based on the individual coding of the focus of the responses, those responses that focused on information gathering or evacuation, were more likely to engage in preparatory actions.

The second research question focused on the relationships between the computer-coded variables on preparatory actions. Independent-sample t-tests were run on the computer coded variables and the results can be seen in **Table 1** as well. Following the same analysis as conducted for **RQ1**, the only significant finding here was that newspapers were negatively associated with preparatory actions ($M=32.91$ when present, versus $M=38.66$ when absent). Of note, because computer coding is manifest content, we could not code information focus using the computer, because the manifest content for "information" was used in a different coding scheme, which would cause a confound. In the human coding process, the information focus variable utilized the actual information sources as a basis for presence, as well as an inference of a focus on gathering information. In the computer coding process, a schema was developed to include manifest content that was

Table 1. Independent samples t-test for human and computer coding on preparatory actions

	Mean (SD)-HC	t-HC	Mean (SD)-CC	t-CC
TV		-2.526*		-0.138
Present	38.97 (7.93)		38.61 (8.26)	
Absent	36.72 (8.75)		38.51 (7.99)	
Radio		-0.811		-1.182
Present	39.18 (7.54)		39.48 (8.25)	
Absent	38.44 (8.24)		38.38 (7.39)	
Internet		1.449		1.803
Present	37.97 (8.02)		37.75 (8.23)	
Absent	38.98 (8.25)		39.03 (8.04)	
Social media		-1.294		-1.004
Present	39.63 (7.16)		39.45 (7.24)	
Absent	38.38 (8.27)		38.43 (8.25)	
Friends & family		-1.272		-0.583
Present	39.19 (7.39)		38.88 (7.93)	
Absent	38.29 (8.43)		38.44 (8.20)	
App		-1.543		-0.294
Present	39.83 (7.98)		38.77 (7.83)	
Absent	38.34 (8.14)		38.51 (8.20)	
Newspaper		2.252*		2.340*
Present	31.71 (6.27)		32.91 (8.80)	
Absent	38.65 (8.11)		38.66 (8.08)	
Television first		-1.904		0.033
Present	38.97 (7.94)		38.55 (8.41)	
Absent	37.51 (8.51)		38.57 (7.90)	
Internet first		1.130		0.937
Present	37.15 (8.35)		37.68 (8.42)	
Absent	38.68 (8.10)		38.68 (8.08)	
Evacuation focus		-2.880**		-1.551
Present	41.81 (4.31)		40.21 (7.38)	
Absent	38.27 (8.33)		38.39 (8.19)	
Information focus		-2.159*		
Present	38.80 (7.84)			
Absent	34.14 (11.49)			
Preparation focus		-1.269		-0.380
Present	39.43 (7.04)		38.92 (7.94)	
Absent	38.33 (8.34)		38.52 (8.15)	
Safety focus		-1.636		-1.738
Present	40.80 (4.26)		43.22 (8.95)	
Absent	38.52 (8.17)		38.49 (8.10)	

*p<0.05; **p<0.001

more than just the presence of information sources. Based on this analysis, it would appear that computer coding was less explanatory about influences on preparatory actions than individual coding.

Finally, our third research question wanted to compare the human coding and computer coding. Chi square tests were conducted on the variables. As noted in **Table 2**, each variable comparison was significantly different, with nearly all of the human coding finding more instances of presence than the computer or manifest coding. The two exceptions were internet appearing first and the presence of newspapers, which were significant in the other direction. We believe this occurred because of phrasing in both instances. In manifest coding, a computer could not distinguish if the individual was referring to news organizations online or internet apps, so it likely was counted more than humans in these cases.

Also, in the case of manifest coding of apps, the first time we ran the data (apps), we found very significant differences between human and computer coding. After some review, we learned that the computer was counting the letters "app" in other words, such as "approach" or "application." Once we recoded to focus on "app" or "apps," which is noted in the revised apps test, we did not find significant differences.

Table 2. Chi-square tests comparing human and computer coding results

Variable	CC Only (N)	HC only (N)	Pearson χ^2	p-value
Television	0.4% (2)	31.6 % (177)	117.09	>0.001* ^a
Radio	1.1% (5)	1.6% (9)	463.79	>0.001*
Internet	1.3% (7)	5.4% (30)	418.24	>0.001*
Social media	0.2% (1)	1.6% (9)	481.43	>0.001* ^a
Friends & family	4.3% (24)	7.1% (40)	295.19	>0.001*
App	5.7% (32)	0.7% (4)	348.20	>0.001* ^a
App (revised)	10.6% (59)	12.3% (69)	1.25	0.264
Newspaper	1.1% (6)	0.4% (2)	177.62	>0.001* ^a
Television first	1.1% (6)	27.9% (159)	154.43	>0.001*
Internet first	5.6% (32)	1.2% (7)	219.23	>0.001*

A: Fisher's exact test because one or more cells were under 5

DISCUSSION

Our study had dual purposes for this analysis. We employed a content analysis to determine how individuals in the path of Hurricane Michael described the informational sources they sought out to prepare for the storm. Findings indicated that television appeared to be the most frequently employed information source when individuals were preparing for the story. Newspapers had a negative influence on preparatory actions and those individuals who describing being focused on their safety or on an intent to evacuate were more likely to prepare than those who were less focused on safety or evacuation. Our second purpose focused on the differences in manifest content between that which was coded through a computer program versus human coders. In nearly all circumstances, human coding results found significantly more occurrences of media messages than computer coding. Implications of these findings will be discussed below.

Perhaps most enlightening about these findings is the significant differences found between the coded content. Programming Excel to look for keywords that fit television, internet, social media and radio—and various slang terms—still does not get at the nuances that human coding can capture. Even among fairly manifest content, differences abound, as has been seen in prior studies (e.g., Guo et al., 2016; Lewis et al., 2013). The reasons for this finding likely stem in part from slang terminology, and for example, differences in the way television and digital video were interpreted by respondents. Although we used several terms to account for television, it's clear that human coders interpreted more instances of television being used as a source than computer coding. Perhaps then, it is time to re-conceptualize and operationalize television? Is a streaming video considered television? That may well be where some of the differences are, although future research would need to more clearly investigate it.

More importantly, however, is the support for the idea that simple manifest coding through frequencies counted by computer programs begets mistakes and should be employed with caution. Computer algorithms cannot get into individuals' brains and determine their thinking—one can argue that human coders cannot either—but the level of inference needed to interpret the results may well be more than a computer algorithm can do. Lewis et al. (2013) calls this "contextual sensitivity" (p. 48). At a minimum, studies using computer content as the main data collection tool, should acknowledge the potential validity concerns.

That is not to say the computer coding—particularly for sentiment and images—lacks validity entirely. Look at Alam et al. (2018) and Su et al. (2016) who have proposed hybrid models of human-based computer-aided content analysis. The idea here is that humans are still involved in the coding process, particularly to evaluate ambiguous or unclear text material. Where validity becomes a concern is when frequency counts are solely within the computer, with no review of the ambiguous material. What our study suggests is that without a level of a review, even face validity of the material can and should be suspect. At the other end of the spectrum, human coding has reliability concerns, which is why the need for the statistical tests (Krippendorff's alpha or Scott's pi, for example) is crucial. While not infallible by any means, it does demonstrate a level of verification, which is missing in the computer coding scheme.

As Riffe et al. (2014) suggest, content analysis should focus on manifest content—that is, what can be seen. Latent content should be part of the interpretation. What our study suggests, however, is that significant differences exist in simple manifest content. That is the key takeaway. Certainly, one can argue that a different

coding scheme in Excel would have produced varied results, but that is not the point. In research, particularly large data or rapid research, the data collection pull and subsequent coding can be an afterthought. In effect, a researcher may just pull the data and accept the software programs results with little review. If researchers assume strong validity based on manifest content only, the results could be problematic. We recommend that this should be addressed in limitations of computer-only coding.

Aside from validity, when weighing the pros and cons of human versus computer coding, reliability within computer coding is not going to be an issue. But we would make an argument that efficiency is not as clearly one-sided, as to address the issue of even a hybrid model of human-aided computer-assisted content analysis (e.g., Su et al., 2016) requires some level of individual coding. Thus, computer coding is not as efficient as it would appear at first blush. Nearly all content analyses software programs brand themselves as qualitative content/textual analyses programs, which is one of the reasons this study used Excel, as a pure frequency counting program.

This comparison has additional merit when considering the extension of this coding project into the larger arena of social science research. While mass communication scholars claim to have the most intricate coding schemes for media texts, researchers in other disciplines also engage in quantitative coding. Therefore, the findings of this study have ramifications for both practitioners and scholars who employ computer coding tools as evidence within their own research.

Finally, this study also has some content-related outcomes that will help disaster-related researchers learn more about how best to reach audiences when severe-weather is imminent. It is clear from the results that individuals turn to television—mostly live weather broadcasting—to learn about their situation when a hurricane is approaching. The totals were 82 percent of respondents, according to our human coding results and 50.7 percent of computer coding results. Qualitative data suggests this includes local news broadcasting in extra hours and cable channels such as CNN and The Weather Channel. This indicates that the most effective reach for emergency management officials is likely going to be those broadcast outlets. Newspapers also had a significant negative association on preparatory actions, meaning those who sought out newspapers were less likely to engage in preparatory actions. However, we report these results with caution, as only 1.3 percent of respondents in human coding and 2.0 percent in computer coding results indicated they sought out newspapers.

The other area of note was in the focus of the respondents in their answers. For those who were focused on evacuation (likely those near the proposed eye of the storm) and those who were focused on gathering enough information to make a decision, they were more likely to engage in preparatory actions. Clearly, this finding makes sense, but it deserves more attention, when considering the implications for scholars and practitioners. Scholars who study the Risk Information Seeking and Processing (RISP) (Griffin et al., 1999; Griffin et al., 2004) will note some support from these findings in that those who were focused on information gathering were engaging in preparatory actions, which suggests some level of information sufficiency was needed. Further, disaster researchers could take heart in the finding that respondents who were focused on evacuation were also engaging in preparatory actions. While those theoretical constructs were not the primary point of this study, the evidentiary support is noteworthy.

Limitations

As with all research, this study is not without its limitations. First, this study focuses on content related to one severe weather event, so generalizations to other storms should be conducted with caution. This study used Excel for its computer coding program. Other more sophisticated content and textual analyses programs may have capabilities beyond simple frequencies, which could allow for more direct comparisons in certain areas. Finally, it is important to note that this analysis focused on an open-ended answer to one question within a larger survey. Analyzing content in a more focused study may yield different results.

Conclusions and Future Directions

This study has significant findings that impact both severe weather research and methodological advances in content analysis. Future work may wish to build on the work to determine whether more sophisticated programs could give more insight. Further scholars and practitioners will benefit from the support for a hybrid model of content analysis, along with support for the impact of broadcast meteorology on disaster

preparedness research. Lastly, the singular focus on broadcast media during this severe weather event should help emergency managers focus their attention for media messaging.

Author contributions: All authors were involved in concept, design, collection of data, interpretation, writing, and critically revising the article. All authors approve final version of the article.

Funding: This manuscript was supported by the U.S. Department of Commerce's National Oceanic and Atmospheric Administration under NOAA Award NA14OAR4170098, the Mississippi-Alabama Sea Grant Consortium, the National Hazards Center and the University of Alabama. The views expressed herein do not necessarily reflect the views of any of those organizations.

Declaration of interest: Authors declare no competing interest.

Data availability: Data generated or analysed during this study are available from the authors on request.

REFERENCES

- Alam, F., Ofli, F., & Imran, M. (2018). Processing social media images by combining human and machine computing during crises. *International Journal of Human-Computer Interaction*, 34(4), 311-327. <https://doi.org/10.1080/10447318.2018.1427831>
- Armstrong, C. L., & Towery, N. (2021). Before and after the storm: How individuals hypothetically and realistically respond to media messages about severe weather. *International Journal of Disaster Response and Emergency Management*, 4(1), 46-62. <https://doi.org/10.4018/IJDREM.2021010104>
- Armstrong, C. L., Hou, J., & Towery, N. (2020). The 'Michael' effect: Risk perception and behavioral intentions through varying lenses. *Journal of Extreme Events*, 1&2, 2050007-1-22. <https://doi.org/10.1142/S2345737620500074>
- Ash, K. D., Schumann, R. L., & Bowser, G. C. (2014). Tornado warning trade-offs: Evaluating choices for visually communicating risk. *Weather, Climate, and Society*, 6(1), 104-118. <https://doi.org/10.1175/wcas-d-13-00021.1>
- Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies*, 9(2), 287-310.
- Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit. *Digital Journalism*, 4(1), 8-23, <https://doi.org/10.1080/21670811.2015.1096598>
- Daniels, G. L., & Loggins, G. M. (2007). Conceptualizing continuous coverage: A strategic model for wall-to-wall local television weather broadcasts. *Journal of Applied Communication Research*, 35(1), 48-66. <https://doi.org/10.1080/00909880601065680>
- Griffin, R. J., Dunwoody, S., & Neuwirth, K. (1999). Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. *Environmental Research*, 80, S230-S245. <https://doi.org/10.1006/enrs.1998.3940>
- Griffin, R. J., Neuwirth, K., Dunwoody, S., & Giese, J. (2004). Information sufficiency and risk communication. *Media Psychology*, 6, 23-61. https://doi.org/10.1207/s1523785xmep0601_2
- Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling. *Journalism & Mass Communication Quarterly*, 93(2), 332-359. <https://doi.org/10.1177/1077699016639231>
- Kang, J. E., Lindell, M. K., & Prater, C. S. (2007). Hurricane evacuation expectations and actual behavior in Hurricane Lili. *Journal of Applied Social Psychology*, 37(4), 887-903. <https://doi.org/10.1111/j.1559-1816.2007.00191.x>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. SAGE.
- Landsea, C. W., & Franklin, J. L. (2013). Atlantic Hurricane database uncertainty and presentation of a new database format. *Monthly Weather Review*, 141(10), 3576-3592. <https://doi.org/10.1175/mwr-d-12-00254.1>
- Landwehr, P. M., Wei, W., Kowalchuck, M., & Carley, K. M. (2016). Using tweets to support disaster planning, warning and response. *Safety Science*, 90, 33-47. <https://doi.org/10.1016/j.ssci.2016.04.012>
- Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34-52. <https://doi.org/10.1080/08838151.2012.761702>
- Linardi, S. (2016). Peer coordination and communication following disaster warnings: An experimental framework. *Safety Science*, 90, 24-32. <https://doi.org/10.1016/j.ssci.2016.03.017>

- Mileti, D. S., & O'Brien, P. W. (1992). Warnings during disaster: Normalizing communicated risk. *Social Problems*, 39(1), 40-57. <https://doi.org/10.2307/3096912>
- Mileti, D. S., & Sorensen, J. H. (1987). Determinants of organizational-effectiveness in responding to low probability catastrophic events. *Columbia Journal of World Business*, 22(1), 13-21.
- Moore, T. W., & Dixon, R. W. (2011). Tropical cyclone-tornado casualties. *Natural Hazards*, 61(2), 621-634. <https://doi.org/10.1007/s11069-011-0050-z>
- Morss, R. E., Demuth, J. L., Lazo, J. K., Dickinson, K., Lazarus, H., & Morrow, B. H. (2016). Understanding public hurricane evacuation decisions and responses to forecast and warning messages. *Weather Forecasting*, 31, 395-417. <https://doi.org/10.1175/WAF-D-15-0066.1>
- Riffe, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages*. Routledge. <https://doi.org/10.4324/9780203551691>
- Sattler, D. N., Kaiser, C. F., & Hittner, J. B. (2000). Disaster preparedness: Relationships among prior experience, personal characteristics, and distress. *Journal of Applied Social Psychology*, 30(7), 1396-1420. <https://doi.org/10.1111/j.1559-1816.2000.tb02527.x>
- Sheldon, P., & Antony, M. G. (2018). Sharing emergency alerts on a college campus: How gender and technology matter. *Southern Communication Journal*, 83(3), 167-178. <https://doi.org/10.1080/1041794x.2018.1437467>
- Su, L. Y., Cacciato, M. A., Liang, X., Broussard, D., Scheufele, D. A., & Xenos, M. A. (2016). Analyzing public sentiments online: Combining human- and computer-based content analysis. *Information, Communication & Society*, 20(3), 406-427. <https://doi.org/10.1080/1369118X.2016.1182197>
- Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox-implications for governance and communication of natural hazards. *Risk Analysis*, 33(6), 1049-1065. <https://doi.org/10.1111/j.1539-6924.2012.01942.x>
- Wu, H., Lindell, M. K., Prater, C. S., & Samuelson, C. D. (2014). Effects of track and threat information on judgments of hurricane strike probability. *Risk Analysis*, 34(6), 1025-1039. <https://doi.org/10.1111/risa.12128>
- Zahran, S., Tavani, D., & Weiler, S. (2013). Daily variation in natural disaster casualties: Information flows, safety, and opportunity costs in tornado versus hurricane strikes. *Risk Analysis*, 33(7), 1265-1280. <https://doi.org/10.1111/j.1539-6924.2012.01920.x>

