HOW DAILY JOURNALISTS VERIFY NUMBERS AND STATISTICS IN NEWS STORIES: TOWARDS A THEORY

By

Anthony Van Witsen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Information and Media—Doctor of Philosophy
Environmental Science and Policy—Dual Major

2020

ABSTRACT

HOW DAILY JOURNALISTS VERIFY NUMBERS AND STATISTICS IN NEWS STORIES: TOWARDS A THEORY

By

Anthony Van Witsen

Statistics are widely acknowledged as an essential part of journalism. Yet despite repeated investigations showing that routine news coverage involving statistics leaves much to be desired, scholarship has failed to produce an adequate theoretical understanding of how statistics are employed in journalism. Earlier research showed many journalists think anything counted or measured and expressed in numbers represents a form of unarguable truth, which may affect whether they think statistical information should be checked or verified. This study examines the verification process in detail by combining 1) qualitative interviews with fifteen working journalists about their attitudes, decision making and work practices regarding statistics; 2) an analysis of manifest statistical content in a sample of the stories created by these subjects; 3) an item-by-item examination of the decision-making processes behind each statistic in each of the sampled stories. Based on the results, I conclude the subjects did not have a single standard for verification, but followed a range of practices from simple reliance on authority at one end to careful examination of the methods behind a quantified fact claim at the other. Theoretical reasons for this are explored.

ACKNOWLEDGEMENTS

This research could not have been accomplished without the generous help of the fifteen working journalists who took large chunks of time out of their busy schedules to answer my many persistent and sometimes intrusive questions about how they did their day-to-day work. The same is true of the eighteen journalists (and one student) whose talks with me in 2017 formed the basis for earlier research that this work built on. Though I can't name any of them, they have my sincere gratitude for tolerating my nosiness. When it comes to tough questioning, you all showed you could take it as well as dish it out.

Fortunately, I can name the members of my doctoral committee: Profs. Manuel Chavez, Eric Freedman, Stephanie Steinhardt and committee chair Bruno Takahashi, whom I want to thank for insightful readings and feedback at all stages of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

It is widely agreed that numbers and statistics are an essential part of journalism (Curtin and Maier, 2001; Harrison, 2016; McConway, 2016). The origins of this state of affairs are harder to agree on. Some reasons might seem as straightforward as the fact that so much news is measured news: surveys, economic statistics, vote totals as well as measures of the social problems, such as crime rates, on which much of journalism focuses. Cohn and Cope (2012) said:

> Even when we journalists say that we are dealing in facts and ideas, much of what we report is based on numbers. Politics comes down to votes. Dollar figures dominate business and government news…. Numbers are at the heart of crime rates, nutritional advice, unemployment reports, weather forecasts, and much more. … The very way in which we journalists tell our readers and viewers about a medical, environmental or other controversy can affect the outcome (p.3).

Behind this preoccupation with numbers is the widely held belief that statistics are objective and scientific, a way of comprehending reality that transcends the difficulties of individual opinion or perception (Anderson, 2018; Porter, 1996; Strathern, 2000). Statistics, in other words, are important because they align with and support journalistic ideals of objectivity and separation of fact from opinion; therefore, mistaken or cursory use of them compromises these goals. While this belief may be true, it may also be too simple (Alonso and Starr, 1987). Political scientists have repeatedly shown that statistics on population (Conk, 1989) income (Jencks, 1989), ethnicity (Petersen, 1989) or more controversial matters such as human sex trafficking (Merry, 2016; Warren, 2010), are less transparent than they seem. This is partly because, as many researchers have recognized (Alonso and Starr, 1987; Brandão and Nguyen, 2017; Lugo-Ocando

and Brandão, 2016; Lugo-Ocando and Lawson, 2017) statistics can never be completely separated from the way the things they measure are conceptualized. However it also sometimes happens intentionally (Lugo-Ocando and Lawson, 2017) when agencies that create statistics tailor the definitions of concepts so as to produce numbers that encourage agency policies or efforts to gain attention. That is, the use of statistics for larger needs (gaining media attention, legitimizing policy, seeking political support) is not something that takes place after the "raw" numbers are created but is sometimes integral to their construction. In that sense, the question of whether statistics are honestly compiled and reported has multiple dimensions and touches on deeper questions than simply matters of incompetence or fraud.

Journalism researchers have begun to raise questions about how these issues apply to the news. What do journalists know about the contingent status of the numbers they try to report faithfully every day? Where do they get numbers from and how do they decide which ones to trust? When do they see numbers as problematic and when not? For example, Lugo-Ocando and Lawson's (2017) study about news coverage of poverty suggests that the very concept of poverty has always been contested and that contemporary statistics about it are inseparable from modern ideas of economic development. This definition began in the first developed economies, which tend to treat their own history as containing the only possible meaning of wealth and poverty. It is these definitions that shape concepts of who can speak about poverty or who can define its "real" meaning as well as "real" standards for measuring it. While these internal disputes may not be covered in the news, their effects still appear in different ways of operationalizing poverty for measurement purposes, such as absolute vs relative deprivation. Unable to produce alternative statistics themselves or interrogate or analyze the poverty models behind them,

journalists tend to report poverty statistics not as an interpretation, but as raw facts that serve as the basis for subsequent interpretation and policy choices.

Similarly, Lugo-Ocando and Brandão (2016) showed how misuse or misunderstanding of widely accepted statistics about knife-related crimes, such as stabbings, can encourage news consumers to misunderstand the scope and nature of a problem. Crime statistics, even when reported accurately, are partly a measure of deviations from accepted norms, expressed in laws which must be defined by some individual or group, whose purposes or values may not be clear and whose methods of enumeration rarely attract widespread attention. The authors concluded that most official crime figures reinforce views of crime derived from law enforcement, prosecutors and politicians rather than those of others such as offenders themselves, social workers, or academics, and reflect the biases and methodological limits of their creators. This reliance on official statistics and the concepts behind them is true even for numbers that are relatively uncontroversial such as scientific statistics which might be thought free of the so-called politics of numbers (Alonso and Starr, 1987; Brandão and Nguyen, 2017). The issue is even more pronounced for phenomena that are new, controversial or concealed, such as drug or sex trafficking. Debates over measuring these can lead to heated battles over how to define or even detect the existence of the subject of measurement (Andreas and Greenhill, 2010; Merry, 2016; Weitzer, 2007).

Because statistics are almost always the product of exactly the expert and authoritative sources on whom journalists rely (Fishman, 1980; Gans, 2004; Tuchman, 1972), the gathering and reporting of the figures that derive from these processes is embedded in journalistic routine, strongly influencing what ends up in finished stories. News judgment in practice is sometimes treated by journalists as “common sense” although both Schudson (1989) and Tuchman (1972)

showed that it is also grounded in tacit assumptions about reality. Carey (1988), for example, says the various forms of news writing are not transparent "facts" but a narrative form whose rhetorical effects exist partly to convey a feeling of neutrality and transparency. This is even true of the inverted pyramid story structure with the five-W "who-what-where-when-why" lead paragraph widely used in "straight" reporting. The assumed neutrality and transparency of numbers may be useful in journalism partly because it contributes to the effect of facts speaking for themselves.

A better understanding of how journalists handle statistics can offer insight into what the media habitually treat as unassailable, inarguable fact. If journalists believe anything expressed in numbers closes off debate, this belief may affect whether statistics are subject to cross-checking and verification. The issue is even more important for journalists who regularly cover science or the environment, whose role almost mandates that they deal with measured knowledge claims, even more so than their colleagues. Science and environmental journalists run into the full range of problems with numbers including issues of risk evaluation or perception, poor problem definition, and uncertain outcomes. Their assignments require them to play multiple roles including translating complicated intellectually precise matters for audiences that may not grasp them easily. They may or may not be better at these tasks than their non science and environmental colleagues; the literature is ambiguous (Bell, 1994; Crow and Stevens, 2012; Dunwoody, 2004; Giannoulis, Botetzagias, and Skanavis, 2010; Gibson et al, 2016; Mcinerney, Bird, and Nucci, 2004; Vestergård, 2011; Wilson, 2000, 2002). For all these reasons science and environmental journalists should benefit more from better understanding of statistics construction, making questions about what they know about statistics potentially more immediate and relevant.

This research examines the verification of statistics in detail, combining an analysis of manifest statistical content in news stories about science or the environment with an item-by-item examination of the thinking processes of the journalists who created them. The mixed-methods approach was chosen in order to understand a complex phenomenon that can operate at many levels at once and doesn't easily lend itself to variables that can be operationalized and counted. Analysis focuses on decision making about sources, trust, source checking and statistics placement. It extends previous work about statistics (Brandão and Nguyen, 2017; Lugo-Ocando and Brandão, 2016; Lugo-Ocando and Lawson, 2017), which focused on broad conceptual problems with statistics construction in journalism without typically getting down to cases. It also adds to research on journalistic verification (Barnoy and Reich, 2019; Diekerhof and Bakker. 2012; Godler and Reich, 2017, 2017b), which studied the choices journalists made in individual stories without looking at the choices they made about statistics as a particular kind of factual claim. This research aims to connect the two, showing links among content, motive and decision making with regard to verification of numbers, while also answering questions about what decisions each journalist made and what they were trying to accomplish with each number included in the story.

RELEVANCE OF THE STUDY TO CONTEMPORARY NEWSGATHERING

The problems this research addresses were spotlighted recently when the coronavirus pandemic not only put scientific expertise "in high demand" (Scheufele et al, 2020, n.p.) but also put a premium on innovative ways of reporting that expertise. In these unusual, perhaps unique circumstances, a few journalists "opened up the black box" and produced reporting that brought readers detailed information about how epidemiological models are created (Foley, 2020; Tufekci, 2020; Yong, 2020) as well as about the challenges involved in counting fatalities (McCarthy, 2020). This matters because journalists who have a critical understanding of how statistics are conceptualized and constructed are less likely to place undue and misleading emphasis on which models and numbers are "right" or "wrong."

However limited, this kind of reporting on the concepts and construction behind statistics represents a step outside what scholars (Fishman, 1980; Gans, 2004; Tuchman, 1973) have repeatedly found about how journalists do their jobs. It is tempting to attribute this to changes in newsroom norms brought on by the crisis environment of the pandemic. However as far back as 2016, reporters at Pro Publica questioned the statistical programs used by some courts to create sentencing guidelines for people convicted after trial (Angwin, Larson, Mattu and Kirchner, 2016). Originally created to make criminal sentencing fairer and more objective, these programs used an algorithm based on past crimes, education and employment to predict their chances of committing more crimes in the future. Although the intent was to safeguard criminal sentencing from the whims of judges, Pro Publica's reporting showed that one algorithm, called COMPAS, had about 60 per cent accuracy, only a little better than random, and had a built-in bias against racial minorities.

Clearly, this story, with its social justice implications, stemmed from the authors' recognition that it is possible to question quantified truths that seem superficially unarguable. But even this reporting is related to contemporaneous research questioning the contested concept of "big data," particularly the idea that big data "offer a higher form of intelligence and knowledge that can generate insights that were previously impossible," imbuing them "with the aura of truth, objectivity, and accuracy" (Lewis and Westlund, 2015, p. 449). Some of this research closely parallels similar work in science and technology studies. Treating scientific objectivity as something that could be studied in itself, scholars such as Latour (1987; 2013), Hacking (1990; 2006), Harding (1986; 2016) and Kuhn (1963; 1987; 2012) investigated the different social, psychological and historical circumstances that helped lead to the creation of particular kinds of scientific knowledge and how some of that knowledge might have been conceptualized differently under different conditions. Although conducted outside the communication research discipline, these findings are directly relevant to questions of how and why journalists believe in numbers and decide which numbers to trust.

It seems unlikely that the journalists who pioneered new ways of examining data were reading the scholars; far more probable that both groups were responding to a changed climate of opinion in which data's penetration into every cranny of life engendered a deeper examination of its underlying premises. A hopeful perspective is that this combination of events will become a "teachable moment" in which more journalists learn about the existence of alternative ways to report the strengths and shortcomings of quantified knowledge claims. This might make them less likely to pass along mis- and disinformation. At the very least, the difficult challenges of reporting the pandemic give increased urgency to Donsbach's call (2014) for all journalists to acquire systematic and formal knowledge of the topics they cover. Donsbach, along with Nisbet

and Fahy (2015) believed that so-called "knowledge-based journalism," especially in science and environmental reporting, might help reverse journalism's declining credibility and influence and reclaim traditional journalistic objectivity from the distortions and exaggerations of bloggers and advocates.

These lessons are also important because the problems caused by uninformed ideas about statistics and data are likely to increase in the future, not only because of the growth of algorithms and big data in contemporary life generally, but also the increased use of data-based reporting, what Meyer (2002) called "precision journalism." With numbers and data increasingly embedded in every decision, every barcode, mag strip and smartphone, with computation and data reporting increasingly common in newsrooms, the tensions between what statistics seem to be and what they are will become increasingly salient for journalists, possibly inescapable. The problem is not only that seemingly neutral statistics always have their origins in particular definitions of phenomena. It is also that numerical representations of things shape perception, frequently in ways that are all the more profound because they are prerational (Tooze, 2001).

The remainder of this dissertation will proceed as follows: it will begin with a literature review that covers the current state of knowledge on the origin and construction of statistics, the cultural role played by statistics, including their impact on journalism, as well as a summary of how previous research on journalism's handling statistics has tended to focus on shortcomings by individual journalists while, until recently, neglecting broader issues of how statistics are generated and conceptualized. It will then link these issues to questions of how statistics in journalism are verified or not and review current scholarship about journalistic verification, particularly the concept of "evidence of evidence" first advanced by Godler and Reich (2017).

Evidence of evidence means journalists treat their existing knowledge about a source as a sign that the source possesses substantive knowledge as well, making this a form of verification itself.

From these, the dissertation develops research questions that use methods developed in verification research to study how journalists' beliefs about the special characteristics of statistics might shape their verification choices. It next describes the methods used to answer these questions, including semistructured qualitative interviews about journalists' thinking with regard to statistics and a case-by-case structured qualitative interviews to determine how particular statistics were verified. Results of the two methods are presented to show where they align or fail to align with previous research. The subsequent discussion considers how the new findings do not disconfirm previous research, but complicate it, showing some of the previously unrecognized ways journalists locate evidence to verify statistics, along with the times and places where they use evidence of evidence about statistics as a substitute for evidence itself, particularly where a single agency has a monopoly on statistics production. It advances some ideas about the mechanisms and processes that keep epistemic standards with regard to statistics are so flexible, what schema this follows, and the purposes it serves. It then follows this with recommendations for future research (including a research proposal to test these ideas) and some practical ways to improve journalists' comprehension of and handling of statistics.

LITERATURE REVIEW

Statistics have been widely studied as a set of analytical tools. However, statistics as social phenomena in themselves with a history, sociology and politics, as well as quantification as a conceptual category of knowledge, have never been brought together in a single discipline and are available only in partial form through multiple disciplines and academic fields including communication research, political science, rhetoric, sociology, science and technology studies and anthropology. Before turning to the issues in journalism related to the use of statistics, I first discuss how numbers and statistics are perceived from a societal perspective. Accordingly, this review of the literature summarizes the main findings from these disciplines as they bear both on statistics and on journalists' use of them.

The origins and construction of statistics

Study of the conceptual, social and psychological factors involved in statistics production originated in political science (Alonso and Starr, 1987; Amberg and Hall, 2010, Andreas and Greenhill, 2010), demographics (Prewitt, 2013), history (Porter, 1986, 1996), rhetoric (Fahnestock, 1986) Sociology (Espeland and Stevens, 1998, 2008) and anthropology (Boellstorff, 2013; Merry, 2016; Strathern, 2000) with different disciplines investigating different dimensions of it. Political scientists such as Alonso and Starr (1987) say all acts of quantification involve value judgments about what to measure and how to operationalize it, the boundaries between categories and between data and noise, along with how to carry out the counting process, that, collectively, derive from presuppositions about the object being measured. In this sense, all statistical activity is politicized, regardless of the quality and professionalism of the measurement process.

One example is Merry's demonstration (2016) of how the official Trafficking In Persons statistics (TIP), created by the U.S. Department of State grow out of a perspective that considers sex trafficking solely as a matter of exploitation of powerless victims by criminals. In Merry's account, the concept of human trafficking and international migration as a whole is a complicated and varied process; not everyone who moves across international borders for purposes of commercial sex is an innocent victim of coercion. Some are coerced, some do it voluntarily, some so-called victims do not feel themselves to be such and resist the efforts of reformers to control them. Yet according to Weitzer (2007) and others, the State Department began creating the annual numbers under pressure from a coalition of conservative and feminist groups who believe the entire sex industry should be abolished because it exploits women. Others, including different sectors of the women's movement, believe commercial sex is a legitimate economic choice for some women. But it was the anti-trafficking forces who persuaded the U.S. government to get behind their definition of the problem to create the annual TIP statistics.

In a similar spirit, other researchers have shown how different forms of knowledge production are dependent on particular concepts of that knowledge. For example Steinhardt (2019) discussed creation of scientific knowledge in which concepts of data grow out of particular ways of deciding, for example, which categories of people in medical research are defined as subjects of study and which are categorized as unclassifiable. Fioramonti (2013) showed how gross domestic product statistics grow out of a model of the economy that favors industrial production while neglecting negatives such as pollution or unmeasured labor such as housework, shaping public understanding of the meaning of economic progress. Lugo Ocando (2017) discussed how

some crime statistics grow out of reports of actual incidents while statistics for minor and unreported crimes are based on surveys.

Another related current of research in history, philosophy, and science and technology studies investigated statistics as part of a larger effort to understand how different forms of knowledge are related to the specific circumstances in which each form is created.

For example, Casper and Clarke (1998) analyzed the social practices by which different knowledge-producing technologies (such as the Pap smear for cancer screening) are used in order to understand the assumptions behind them. Like Lewis and Westlund, Boyd and Crawford (2012) criticized the idea of quantification as the embodiment of objective truth, particularly the belief that with enough data, numbers can speak for themselves. These authors cited Gitelman's (2013) insight that data must be conceptualized or imagined as data in the first place, then cleaned, before it can be analyzed. This always involves processes that are discipline-specific, with an unavoidably subjective element. Other researchers such as Hacking and Hacking (1990) analyzed the historical replacement, over several centuries, of determinism by randomness, both as a concept and as a practice through mathematical rules of probability. This led to the growth of statistical bureaucracies with the power to classify according to categories which frequently did not even exist before the agencies created them. The Hackings call this "making up people." (p. 3) The power of these systems of classification went on to shape peoples' ideas of which things were normal and which deviant, especially in regard to human behavior.

Almost all these researchers agree that the numbers produced thereby have immense influence on public perception of the things they were created to measure. Because these numbers reach the public through the media, news reporting (or misreporting) of statistics almost always plays a role in this influence. When these essential but partly normative numbers are made public, their

numerical expression can make them seem like something beyond norms (Amberg and Hall, 2010; Fahnestock, 1986; Strathern, 2000), shaping public ideas about the size and scope of the things they were created to measure, including the social problems on which much journalism focuses. Advocacy groups and social movements, following this logic, know statistics give them credibility and improve access to news coverage (Best, 1987). These groups sometimes harness the rhetorical power of pure numbers to take into account the needs of specific constituencies including the media (Della Porta and Diani, 2006). In that sense, possession of some kind of statistics is almost a ticket of admission to the public sphere.

Lugo Ocando and Brandão (2016) say statistics in journalism are a commonly accepted language which not only reinforces the picture of reality they generate but also reinforces the use of statistics to produce this reality. The recursiveness is abetted by the fact that the calculations and judgments necessary to produce statistics typically take place before the numbers become visible to the public or to journalists or are confined to footnotes where they rarely attract coverage (Bhatti and Pederson 2015; Prewitt, 1987; Rose 1991). As a result, when statistics appear in the news, they are frequently treated as if the phenomenon being measured were identical to its operationalized indicator. This process, which effectively amounts to conflating the temperature with the thermometer, has been labeled "surrogation" (Choi, Hecht, and Tayler, 2012). It sometimes leads to such phenomena as "teaching to the test" in education (Volante, 2004) or to political decisions to tailor economic statistics to shape perceptions of poverty and wealth (Agren, 2016). While this process frequently emerges via statistics in the news, it also stems from the larger phenomenon of trust in numbers, which has both instrumental and psychological components. Porter (1996) says that this trust, and the promise that statistics can create knowledge independent of the people who create it, makes quantification particularly

appropriate for communication that goes beyond local boundaries. This view of objectivity, of meeting the moral demand for fairness, makes numbers appealing to unelected officials who must justify their decisions on some basis other than winning an election. Officials, in other words, may use numbers in decision making both for the practical value of what they measure and also to avoid the public perception of arbitrariness. Porter's larger point is that even though quantification creates forms of intersubjective knowledge that go beyond particular communities, this so-called "struggle against subjectivity" (p. ix) also has a value that serves narrower political interests. Lugo-Ocando and Lawson (2017) say institutions that create statistics are concerned both with objective measurement and the public support that follows the perception of that objectivity and take both into account in the choices they make. The logic of media needs, in other words, is not an add-on to statistics, but part of their creation.

Numbers as cultural and rhetorical artifacts

Number perception, however, includes not only individual psychological components but larger cultural meanings of numbers. De Santos (2009) and Berman and Milanes-Reyes (2013) both recognized that the meaning of numbers can change when they move from the specialized environment where they were created and reach larger audiences. Berman and Milanes-Reyes (2013) studied the Laffer curve in the U.S. Congress *Congressional Record.* Where Republicans discussed it as a testable hypothesis about the relationship between taxes and government revenues, Democrats treated it with scorn and derision. A notable aspect of this research is the way media played a role in the popularization and transmission of cultural attitudes toward these statistics. This is an example of the hierarchy of influences process in which journalism incorporates ideas from the culture in which it operates (Reese and Shoemaker, 2016), including

numbers that acquire special meaning. However, there is little or no research on the way journalists make decisions about numbers with this cultural power.

Although statistics are often thought of as solely mathematical in nature, the ways they are presented also incorporate multiple rhetorical effects. Bodemer, Meder and Gigerenzer (2014) referenced a breast cancer study in which mortality fell from 5 in 1000 without screening to 4 in 1000 with screening. Depending on presentation format, this can be seen as an improvement of 20% (from 5 to 4), a relative risk reduction, or an improvement of 0.1% (from 5 in 1000 to 4 in 1000), an absolute risk reduction. This is but one example of how changes in the rhetoric of statistical presentation can affect perception of their mathematical meaning.

Examining numbers on their own terms, Bodemer, Meder and Gigerenzer (2014) also found some numbers in newspaper headlines (such as a budget figure of 4 trillion 360 billion spelled out with all the zeroes, stretching across 2/3 of a page) were there to "excite by the iconic element: the unusual length, the repetition of 10 zeroes in a row shows how huge, how outstandingly large is the budget" (p. 358). Looking even beyond the reference value of numbers as substantive measurements, Koetsenruijter (2011) found experimentally that people sometimes respond to numbers simply as signals of precision and truthfulness. The numbers, in other words, functioned as rhetorical symbols separately from their substance (Roeh 1989; Roeh and Feldman 1984). Koetsenruijter named this phenomenon "number paradox" to describe journalism's tendency to load articles with numbers even though surveys show most readers do not remember them. These tendencies originated in the larger culture rather than in journalism, but are frequently incorporated into the presentation of statistics in news stories through the hierarchy of influences model (Reese and Shoemaker, 2016).

Statistics and journalistic decision making

At the individual level, communication scholars such as Ahmad (2016), Brand (2008), Maier (2002, 2003), McConnell (2014) and Moynihan et al. (2000), have documented specific examples of mishandling of statistics in journalism. Reasons offered for these shortcomings differ. While Hand (2009) recognized that statistics contain many nonmathematical elements, some of these researchers continue to blame journalists for lack of math skills. Journalists, in this view, are seen as more comfortable with qualitative thinking, feelings, and words than with mathematical concepts and numbers. Curtin and Maier (2001) found a corollary phenomenon: "math anxiety", that caused journalists to gravitate away from statistics or even state that they became journalists to avoid mathematical thinking. However, Harrison (2016) found that while most journalism students had not taken science courses, he also found the division between narrative and non-narrative knowledge existed largely at the level of anecdote:

> Colleagues at my institution recall asking undergraduate students what they thought of mathematics—many replied that it was a dislike of mathematics at school that led them to choose journalism as a degree subject in the first place. (p. 1)

Harrison believes that when such anecdotes are repeated often enough they metamorphose into a taken-tor-granted truism regardless of their status as testable truth. This makes it easier for students to think of themselves as word- or numbers-oriented. Callison, Gibson and Zillman (2009) reached the same conclusion but also concluded that much is not known about how people think about data or how numerical ability predicts their ability to handle statistics.

These findings show that in addition to their mathematical components, statistics contain many dimensions—social, cultural and political—that are all intrinsically value-laden. While many of these are directly relevant to how numbers are sourced, trusted, checked or written

about in journalism, few of them have been taken up by media researchers. Those who investigated the rhetorical power of numbers (Bodemer Meder and Gigerenzer, 2014; Koetsenruijter, 2009, 2011; Peters, et al. 2006; Roeh and Feldman, 1984) focused on individual psychological perception but paid little attention to the micro-decisions journalists must make when they cull from a large mass of statistics or how they make choices when a topic has both statistical and non-statistical elements. Communication researchers (Curtin and Maier, 2001; Harrison, 2016) focused on cognitive or psychological issues of numeracy, math skills or math-anxiety among journalists without examining other factors, such as professional norms, organizational routines, judgments about audience expectations or cultural influences. Any or all of these can affect what reporters or editors think about the validity of any fact claim at all; however they may not have the same effect on how journalists think about statistics than how they think about other categories of facts more widely recognized as normative, such as those contained in political debates. Van Witsen (2018) found journalists are usually unaware of the constructed nature of most statistics and frequently believe numbers provide access to a kind of truth not available through eyewitness accounts or interviews. It is not clear what leads to this certainty. It could be a response to the perceived certainty of science and numbers as concepts or to the authority of scientific and statistical institutions as widely accepted and trusted sources. Some journalists may not distinguish between the two.

Brandão (2016), studying science news in Brazil and the United Kingdom, inquired into the relationship between statistics and journalists' concepts of objectivity. The study found numbers were frequently used to legitimate the science, used less to convey information than to make news appear objective. This was based on findings showing relatively even distribution of numbers of statistics across hard news stories, while editorials frequently cited one or none. This

led to the conclusion that the use of multiple sources indicates the need to present hard news, verified by statistics, as legitimate. Over 70% of cases contained no information at all about methods or about the kind of study being reported on. This served to link authoritative sources with the concept of objectivity in science and also with journalistic conventions of objectivity and the verification norms and routines found by Barnoy and Reich (2019), Diekerhof and Bakker (2012) and Godler and Reich (2017).

This study builds on two previous studies about how journalists think about and use statistics in daily news. The first (Van Witsen, 2018) showed that many working journalists believe as a rule that statistics are so real as to be unchallengeable. They tend to be aware of the problems of a particular number only when they have learned about its origins through long experience with it. Overall, they follow statistical conventions observed by the beats they cover in determining which numbers to use. This follows theories of trust in news sources and the larger belief in the transparency of measured reality in general, which reaches journalists through the hierarchy of influences process (Reese and Shoemaker, 2016). The second (Van Witsen, 2019) also showed how this certitude emerged in a single story, the 2017 announcement by the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA) concerning record average temperatures for the previous year, 2016. A discourse analysis of 95 news stories showed the heavy use of certainty markers, rhetorical expressions intended to convey the absence of doubt. This was consistent with research showing that journalists relied strongly on authoritative sources to determine what counted as good measurement (Van Witsen, 2018). Numbers produced by a small group of influential scientific sources were always conveyed with rhetoric that either eliminated doubt about them entirely or minimized its importance. If, as the previous two studies indicate, journalists generally believe

statistics are authoritative and undebatable, these beliefs may affect whether they think statistics are worth checking before they go on to publish, broadcast or post them. This is what this study investigates.

Verification of statistics in journalism

Research on verification in journalism has yet to fully answer the broad questions of how journalists verify facts or, when they do, what methods they use. Godler and Reich (2017) argue that philosophically sophisticated truth and certainty are unattainable in journalism even in principle; therefore, verification in news involves many processes beyond formal tests of external truth analogous to scientific methods. Actual newsroom practices show journalists may not even have a single definition of what constitutes an indisputable fact or a single basis for employing systematic, formal tests such as checking an assertion with alternative sources. Overall, Godler and Reich (2017b) agree with other researchers (Fishman, 1980; Gans, 2004; Tuchman, 1973) that journalists rely on a broad set of norms to do their work, including their own judgments about, experience with, and relationships to sources. These manifest themselves in routines such as regular contact with authoritative or official sources of news (such as government officials or large institutions) and routine attention to places thought to reliably generate news (such as courts or political bodies). The trustworthiness of a source may by itself be a basis for determining both that a story exists and also that it is true.

When journalists do seek to verify facts, their rules appear to be largely tacit, in accordance with Schudson (1982; 1989). Diekerhof and Bakker (2012), studying Dutch journalists who wrote in-depth stories on their own initiative (as opposed to stories assigned from above by editors), found that even though their subjects recognized the importance of fact-checking, their

actual practice showed many facts were left unchecked. Their subjects checked sources when they were perceived to have an interest in what the information contained or when checking was easy to do; that is, for reasons that relate to organizational needs. Barnoy and Reich (2019) found the facts most likely to be verified are leaks, exclusives, unplanned events or events perceived as important. They concluded extensive verification occurred in only a minority of stories such as complex stories in which basic facts were in serious dispute or where conflicting sources were perceived to have an interest in different versions of the facts.

In addition to actual cross-checking, journalists must sometimes decide which things they regard as sufficiently well established to require no checking at all and can therefore be used as building blocks to verify larger truths. Godler and Reich (2015; 2017) argued that while journalism could not do metaphysically sophisticated investigations of a truth claim, its accomplishments were not totally subject to authority. How, they asked, can journalists ever hope to meet well-supported standards of what is accepted as true, when they rarely if ever have direct access to knowledge or knowledge-producing processes? They answered this through a concept called "practical skepticism" (Godler and Reich, 2017 p 2) linked to specific practices of knowledge-seeking. Where earlier researchers (Allan, 2004; Ekström, 2002; Ericson, 1998; Ericson et al., 1987; Ettema and Glasser, 1985, 1998; Fishman, 1980; Sigal, 1973) regarded source trust as evidence-free and irrational, Godler and Reich believe existing forms of journalistic fact gathering and verification should count as evidence, if properly understood. In this view, all forms of trust are not created equal and the decision to trust information from a human or institutional source cannot automatically be dismissed as unrelated to evidence. Basing their ideas on the epistemologist Alvin Goldman (Goldman, 2002; 2010), they argued that trust is a rational basis for judging truth when the party doing the trusting has some evidence about the

party being trusted even if this isn't direct evidence about how the fact claim itself was ascertained. This process, which they called "evidence of evidence" or "second order evidence," could extend to such things as sources' past record of accuracy, or willingness to engage critically with their own thought processes and conclusions. This study attempts to empirically document where and when "evidence of evidence" functions, e.g. a journalist's previous history of working with a source or the source's standing in the relevant community that may affect the likelihood of telling the truth.

Statistics have a number of characteristics that might affect journalists' attitude toward checking them, including their reputation as fundamentally neutral and objective (Anderson, 2018; Porter, 1996; Strathern, 2000). Statistics created to measure social issues such as crime or unemployment are frequently taken to symbolize something larger than themselves, such as the success of the various policies they were created to measure. If these characteristics, combined with the cultural and rhetorical power of numbers (Berman and Milanes-Reyes, 2013; Bodemer, Meder and Gigerenzer, 2014; De Santos, 2009; Koetsenruijter 2011; Roeh, 1989; Roeh and Feldman, 1984), lead journalists to see them as beyond argument, this perception may cause them to assign statistics to the category of routine or authoritative information not in need of cross-checking.

While verification researchers have not focused on statistical evidence in particular, many of their methods and conclusions raise are applicable to numbers as a particular kind of news source and knowledge claim. Because journalists are rarely in a position to investigate the truth of numbers themselves, trust is likely to play a larger than normal role in their verification. This suggests an important role for evidence of evidence. In addition, statistics are frequently a product of large institutions, including scientific institutions. Except for journalists doing data-

based reporting (and sometimes for those as well) the information and resource asymmetry between journalists and numbers-creating institutions makes conventional cross-checking impossible. This combination of circumstances may lead some journalists to rely on numbers produced by large government agencies, advocacy groups or scientists, without differentiating among their various products or knowing the assumptions, values and conceptual bases behind their creation. Since these questions are similar to the ones raised by Barnoy and Reich (2019), Diekerhof and Bakker (2012) and Godler and Reich (2017) some of the methods used by these researchers might be used to investigate these issues.

Even though such questions have not been taken up by verification researchers, the existing research is suggestive. Diekerhof and Bakker (2012) found journalists are less likely to check facts when they come from an authoritative source. Barnoy and Reich (2019) found expected events were not verified, particularly when they involved well-understood institutional routines. The authoritative, institutionalized quality of many numbers, such as monthly unemployment figures, might combine with the legitimizing function discovered by Lugo-Ocando and Brandão (2016) to make such news fall into the category of "just the unadorned facts" and therefore not necessary to check. Additionally, the circumstances under which statistics are reported typically leaves little room to pursue them in the depth that might help audiences understand the normative choices that lie behind their construction (nor is it clear that audiences would always have the time or interest). This may serve as a further disincentive to verification.

Existing research on verification tends to focus on specific decisions about specific fact claims by individual journalists, one story at a time, while ignoring the professional environment in which these choices, and the people who make them, are embedded. Many of the verification papers cited above focused on verification decisions by the journalists who wrote the stories.

This sometimes seemed to imply that journalists always initiated verification decisions themselves when in fact that institutional routines and policies may also have played a role. In searching for a fuller understanding of how statistics and verification work together, this study bridges verification research with newsroom sociology, searching for the ways in which organizational expectations influence decisions to verify. Researchers (e.g. Fishman, 1980; Reese, 2016; Stonbely, 2015; Tuchman, 1972) agree that journalists working in newsrooms are continually subject to the pressure of professional norms, routines and expectations and these include accepted ways of determining the accuracy of a fact. Tuchman discusses the standard practice of balancing opposing views of controversial topics in the absence of additional reporting. In general, journalists look to authority, particularly institutionalized authority such as government officials, as sources that can be trusted without checking. Schudson (1982, 1989) among others, argues that the reporter/official connection makes journalism an important tool of official authority, which if true, might have implications for how journalists report official statistics. Sigal (1973) and Usher (2013) found journalists and their editors understand their roles as components of an institutional system quite well and know what the expectations are. The advantages of this (Sigal, 1973) are three: first, it provides general guidelines for what can routinely be regarded as truth to which journalists can refer and which they are not expected to stray far beyond without sufficient justification (such as a senior reporter with superior access or ability to handle sources or data). Second, it permits publishing stories based on a single source as long as it meets the institution's standards for straight news from an authoritative source (e.g. "The President said today…"). Third, it allows journalists to produce major news with a minimum expenditure of resources (Tunstall, 1972). Together, these suggest the traditional newsroom social system may find statistics particularly useful, yet the matter has rarely been

investigated. Verification researchers such as Usher (2013) have recognized that the deep and continuing structures of news work, including time constraints, audience expectations, beat and source limits and decision making hierarchies, have a strong influence over the finished product but have generally not considered how much these influence processes and standards for checking facts.

Schultz (2007) and others believe at least some journalists have considerable capacity to act independently based on their social position of individual actors in the journalistic system. Journalists who possess social capital such as experience, job title, news beat or record of accomplishment and public acclaim (e.g. awards) may have a greater capacity to make individual decisions about when and how to verify; however individual agency can be very difficult to measure. Others such as Cottle (2000) or Stonbely (2015) and Usher (2013) believe early researchers overemphasized organizational dynamics in the newsroom. The presumption of these researchers (Dickinson, 2008) was that the way editorial product is created plays the largest role in what it says to audiences, neglecting the effect of larger economic and cultural forces on news product. One of these forces is the status of numbers in modern culture (Porter, 1996). The overall difficulty with an institutional approach to journalistic decision-making is the tendency of some early researchers to focus on the system alone, as if individual journalists had no agency at all.

Collectively, any or all these forces may affect the way journalists choose what to check and appropriate methods for checking. Some of the organizational factors (Fishman, 1980; Schudson, 1982, 1989; Tuchman, 1972) have already been shown to play a role in verification standards. Because many journalists recognize the authority of numbers, it is likely their verification practices reflect this in some way. However Brandão (2016), one of the few researchers to focus

explicitly on statistics in science journalism, concluded that traditional methods of journalistic verification do not apply in science. Statistics are sometimes used to get attention but in general, journalists do not have any system for using statistics in their stories and tend to present either too few or too many. Brandão says journalists with science backgrounds understand numbers better than general assignment journalists, a finding which contradicts other studies (Bell, 1994; Dunwoody, 2004; McInerney Bird and Nucci, 2004; Vestergård, 2011; Wilson, 2000, 2002). This study seems to clarify the connections among journalism, science, and statistics, particularly in regard to verification.

Verification research and statistics

Because much of the research about statistics discussed above exists in scattered form across multiple disciplines, it has been difficult to bring it to bear on journalism or even at times, recognize its relevance to media research. By bridging research from political science and elsewhere about the conceptual basis behind numbers with communication research on news verification practices, this study seeks to contribute to a better understanding of several related issues about how statistics function in journalism that are increasingly relevant as the datafication of society continues apace. Verification research inquires into what journalists think truth means amid conditions that sometimes compel them to verify with quick, simple means. The perspective of this research stance makes it possible to ask certain kinds of questions about statistics in journalism not otherwise answerable. For example the discovery that journalists believe authorities and experts possess uniquely valuable kinds of knowledge may intersect with their belief in the special epistemic status of statistics to make it less likely that numbers will be checked. Or they might interact in ways that favor certain kinds of checking over others. Past

experience with certain numbers or certain sources might affect the status journalists grant to new numbers that appear related or to new claims by the same sources. The concepts and social arrangements that lead to particular definitions of what counts as data may become more visible to journalists under some circumstances for some kinds of stories than for others.

Combining the two kinds of research orientations referenced above may make it easier to formulate these issues as research questions. Exactly how much about numbers do journalists understand when they report on them in the context of a specific story? How do these vary by type of story, type of media outlet or by the differences between individual journalists? How much interest do journalists show in the conceptual bases behind statistics production and do these affect their verification decisions? As more and more things are datafied, will these concepts become more apparent to more journalists? Does the widespread belief in the special epistemic status of numbers affect journalists' likelihood of verifying them? When journalists do seek to verify statistics, what does that process consist of and how does it differ from verification of other fact claims? How do all these issues function amid the pressures, uncertainties and compromises of daily news production? How do these practices as a whole make (or fail to make) a substantive contribution to finding truths, particular about a form of knowledge production which is growing in importance every day? Verification researchers recognize that despite their discoveries, many aspects of journalistic decision making remain stubbornly contingent and cannot be explained by a simple principle. Nevertheless, patterns exist, which this study seeks to discover.

RESEARCH QUESTIONS

Two previous studies (Van Witsen, 2018, 2019) looked at broad patterns of journalistic thinking about statistics and at how stories involving statistics used language in a way that was consistent with this thinking, without investigating how the stories themselves were actually produced. This study tries to answer that question through mixed methods. It combines semistructured qualitative interviews, an analysis of manifest statistical content in journalists' stories, and structured qualitative interviews about those statistics to shed light on how journalists' beliefs about numbers shape their actual decision-making, including their reliance on particular sources, particular kinds of sources and particular numbers, along with their trust, or lack of trust, in particular sources of numbers, as well as how these are evident in editorial product.

The existing literature shows statistics play many roles based on their origin and construction, and affect people on many levels including as rhetoric and as forms of culture. Because any of these can serve as drivers of journalistic decision making both by individuals and by journalism as a system, the research questions seek to address the relationship between what is known about what journalists believe about statistics and what is known about their methods of verification. It will begin with descriptive questions:

RQ1: Which statistics do journalists consider newsworthy enough to include in their stories?

RQ1a: What are the characteristics of newsworthy statistics (e.g. arising from authoritative sources, regularized or routine sources, or part of newsroom routine)?

Because concepts of what has the status of news cannot be separated from concepts of what doesn't, RQ1b will ask

RQ1b: What are the characteristics of non-newsworthy statistics?

In some cases, journalists regard some information as sufficiently well-established to be considered "evidence of evidence" and therefore not in need of checking; in other cases not. However, the question of which facts require verification may not, by itself, be a separate question from which facts are newsworthy. In other words, journalists may sometimes regard the significance of a new development as contributing to its veracity. But significance itself may derive from previous ideas of a source's familiarity or credibility or its integration into newsroom routines. Leaks or exclusives may be treated differently from more routine stories as well as unexpected or controversial developments. To better understand how journalists make these distinctions with regard to statistics, the study will investigate:

RQ2: What types of statistics do and do not need checking?

Previous research (Van Witsen, 2018) showed that journalists frequently trust numbers-based fact claims and that this trust grows out of both the authority of numbers-creating institutions and the cultural authority of numbers themselves. If this is the case, decisions to cross-check or corroborate numbers may be related to the authority of the source; for example, numbers produced by an advocacy group or an interest group may lead to more frequent efforts to verify than numbers produced by scientists or government agencies. Because authoritative source status is sometimes taken as evidence of the credibility of a statistic, I will inquire:

RQ3: How does the status of a statistical source affect judgments about its credibility? For example, do journalists decide differently for statistics coming from official sources than for statistics originating with NGOs such as advocacy groups? Or do they decide differently for different kinds of statistics?

Organizational norms, expectations, and hierarchies as perceived by journalists play a large role in their decisionmaking. To learn more about how this functions, the study will ask:

RQ4: How do journalists' perceptions of the norms of the organizations within which they are embedded influence individual decisions about checking statistics?

RQ5: What formal tests, if any, do journalists use to check the accuracy of a statistic?

RESEARCH DESIGN AND METHOD: RECONSTRUCTION

Some of the investigators cited above (Diekerhof and Bakker, 2012; Godler and Reich, 2017; Reich and Barnoy, 2016; 2019), who study journalist/source relations have developed a method called reconstruction to better understand how their subjects made decisions about which sources to use, to trust, to check and to publish. These authors argue that journalists' decisions should have manifest consequences that can be learned by observing their thinking or practices. They accomplished this by combining content analysis of a sample of stories with structured qualitative interviews about the specific decisions that went into each choice of a source of information and each decision to verify the information or not.

The strength of reconstruction is its ability to link journalists' principles and beliefs about their work to specific practices (Bruggemann, 2013). Unlike free-floating semistructured interviews (Barnoy and Reich, 2019; Reich and Barnoy, 2016), structured reconstruction interviews can capture the logic behind news priorities, judgments, norms and resources at a detailed level, even zeroing in on a single contact and decision about a single source on one particular story in order to link these with broader principles or practices. Although particular studies may vary, all reconstructions encompass four steps: 1) assembling a population of journalists; 2) sampling their published editorial product; 3) interviewing the creators of each product about each editorial decision involved in its creation; and 4) analysis of the results.

Reconstruction proceeds both horizontally, from one fact to the next across the progress of a story and vertically, through the specific reasons for contacting, trusting or using an individual source of information, and can be documented as such. The issues this process can probe (Godler and Reich, 2017; Reich and Barnoy, 2016) include:

1) Whether journalists make reason-based decisions to include or not include methodology.

2) What reasons were cited in the decision to check or not.

3) Whether checking decisions are based on source type.

4) Whether checking decisions are based on familiarity with a source, importance or authoritativeness of the source, routineness of the information, clarity of the information conveyed, presence of controversy or conflict or organizational norms or policies.

5) Whether the news was exclusive, whether the information was leaked.

6) Whether decision to trust or check was based on formal or informal norms, routines or policies as journalists perceive them.

Other questions may inquire into how the news event's existence was detected in the first place, the role of trust in different types or classes of sources, or how often different types of sources are employed. It is also possible to inquire whether the perceived transparency of statistics engenders trust in itself.

These micro-accounts can then be aggregated into a larger picture, yielding generalizable findings. Because the research does not seek any sensitive nonpublic information, the interviews do not ask for the identity of confidential or background sources. In past studies, reconstruction has emerged as a powerful tool that can study story creation from the perspective or broad principles but can also go down to as fine a level of detail as necessary including a single contact with a single source.

How this study incorporates reconstruction

Despite the strengths of reconstruction methodology, Reich and Barnoy also acknowledge that not giving journalists an opportunity for self-reflection means the method fails to capture the values, beliefs or concepts that lie behind their decisions. This limits the explanatory power of

structured interviews alone. This research, therefore, differs from previous reconstruction studies in two ways: 1) it investigates only decisions about statistics; 2) it will precede the structured interviews with semistructured questions about the subjects' general views of statistics: what they understand data to be, how they think it functions, where they see functionality or dysfunction in data and how they decide which statistics to trust or distrust. Based on Van Witsen (2018), I expect consistent patterns to emerge in the way these subjects think about statistics. The structured interviews will then follow.

The two kinds of interviews make it possible to combine the granular element-by-element aspects of reconstruction interviews with the semistructured element to show how these decisions do or do not relate to how the subjects think about statistics. Semistructured interviews, which are focused but open ended, can yield a richer range of data that captures the individual subjective experience people use to ascertain the meaning of something such as the origins and role of numbers (Hesse-Biber and Leavy, 2010).

Research method components

The full method used to address the RQs had multiple interrelated steps. Specific stages, discussed in more detail below, were 1) locating a population of subjects; 2) locating stories by the subjects and compiling an inventory of all the statistics contained within them; 3) semistructured qualitative interview with each subject; 4) structured qualitative interviews about the source and verification decisions behind each statistic in each subject's story; 5) coding the semistructured interviews; 6) analysis of the structured interviews; 7) repeating steps 1-6 with a second population of subjects; 8) comparing codes in the first population to codes in the second

population to determine whether saturation had been reached. Each of these steps and its theoretical justification is discussed here, beginning with sample size.

Population sample size

In the absence of existing studies of how journalists verify statistics, I used a version of Godler and Reich's (2017) variables but I used a purposive sample instead of a random sample of subjects. According to Merriam (1995), a large sample is useful when variables are well understood, known ahead of time, or operationalized in a recognized way so they can be used to test a priori hypotheses. A small sample, on the other hand, allows for a closer, more detailed understanding of a phenomenon in order to build new theory or when variables are not yet known. For this reason, a small sample was used. If successful, this may later be used to generate hypotheses (Guba and Lincoln, 1994) which can be tested through a large sample. For both large and small sample sizes, ideas of validity and reliability must be related to the purposes and perspective of the study. Interviewing a small number of subjects in depth makes it possible to focus on details of journalistic decision making such as the different factors used to assess the credibility or institutional status of number sources, or what steps are taken to find alternative sources for cross-checking.

Numerous investigators (Brod, Tesler and Christiansen, 2009; Francis et al, 2010; Fusch and Ness, 2015; Guest et al, 2006; Marshall et al, 2013; van Rijnsoever, 2017) have sought a method for determining an adequate sample size for qualitative interview-based studies. Though these researchers do not agree completely, they all focus on achieving saturation, a concept developed for qualitative interviews that represents the hypothetical point at which additional interviews do

not add to theoretical development and the main variations have been identified and incorporated into theory.

Guest et al. (2006) operationalize saturation as the point when additional data adds little or nothing to the development of new codes. In one of their studies, 73% of codes were identified in the first six interviews, 92% in the next six. In a second study, 94% of all codes emerged within the first six interviews, 97% after the first twelve. They conclude most codes emerge quickly in early stages of analysis, then the law of diminishing returns sets in, with fewer and fewer codes emerging from more and more interviews. Given the homogeneity of journalists' thinking about statistics (Van Witsen, 2018), with a consistent division of perspective emerging only between routine reporting and investigative or data-based reporting, it is reasonable to conclude that similar codes will emerge in journalistic decision making. This should be true even for science and environmental reporters given the lack of demographic or professional diversity turned up by previous research (Bell, 1994; Dunwoody, 2004; Mcinerney, Bird, and Nucci, 2004; Vestergård, 2011; Wilson, 2000, 2002).

Francis et al. (2010) suggest that researchers specify a sample size, then specify a priori how many more interviews need to be conducted without any new codes emerging in order to conclude that saturation has been achieved. After the first group of interviews has been coded, the researcher proceeds to interview and code the next group of subjects. Interviewing can stop when there are three consecutive interviews with no new codes. These authors found that even in studies of different types of behavior and different populations, the number of new codes began to plateau after six interviews. They concluded that a minimum sample size of 13 was enough for their study of the theory of planned behavior. Based on these guidelines, I decided to interview six subjects initially and code their responses. I then interviewed an additional six subjects and

coded their responses to see whether new codes emerged in the second round of interviews. When none did, as an additional check, I located and interviewed three more subjects. When no new codes emerged from these interviews, I concluded the interviewing stage of the research.

Study population: why science and environmental journalists were studied

Science and environmental journalists have a special relationship to quantified knowledge. Their professional role requires them to translate scientific knowledge (Dunwoody, 2004), which is frequently arcane and technical, to multiple nonscientific audiences. Because of the importance of statistics in science, they frequently deal with numbers-based findings with high-stakes implications for different policy choices involving issues of risk evaluation or perception, poor problem definition, and uncertain outcomes. Their professional ideals require them to write precisely and sensitively about the interaction between science and its policy aspects for audiences that may not see them as separate things. Bødker and Neverla (2012) say environmental journalism in particular lies at the intersection of science and values. In addition, they frequently deal with advocacy and interest groups who use scientific knowledge or statistics as tools in their own agenda statistics which may not represent the only interpretation of the issue.

A group of journalists that regularly covers high-stakes, policy-relevant, numbers-based science logically ought to benefit from greater understanding of numbers, making them a valuable group to investigate. What does this group know about how statistics are used in science and in policymaking? Given these stakes, it may be surprising that there is little evidence science journalists with advanced degrees or scientific training produce higher quality reporting (Bell, 1994; Dunwoody, 2004; Mcinerney, Bird, and Nucci, 2004; Vestergård, 2011; Wilson, 2000,

2002) than their non trained colleagues. Nor is there much evidence that science journalists have significantly different education or training than non-science journalists (Crow and Stevens 2012). In addition, while different environmental journalists see themselves as playing different roles (Giannoulis, Botetzagias, and Skanavis, 2010; Gibson et al, 2016) including straight reporters of facts, interpreters of facts or mobilizers of public opinion, none of the journalists studied in the above research had advanced degrees in environmental science. Crow and Stevens (2012) found general assignment reporters covering science and environmental stories did not seek out additional science training or regret their lack of it while Giannoulis et al (2010) found environmental journalists in Greece did not think thei8r reporting would improve if they thought more like scientists. A group of journalists with special responsibilities but no special training (and little evidence they would benefit from that training) raises questions about how their thinking about statistics and their methods of verification function. Does their performance on either variable differ in any significant respect from what is already known about their peers?

How the study population was assembled

Because no complete sampling frame of science or environmental journalists exists, the study population represents a convenience sample. Subjects were located through two notices apiece on separate email discussion boards sponsored by the Society of Environmental Journalists (SEJ) and the National Association of Science Writers (NASW). The two NASW notices drew 13 responses; the SEJ notices drew 5, but not all the initial responses were ultimately available for interviews or met the selection criteria. The discussion board notices were augmented by snowball sampling through references provided by initial subjects as well as cold-call email messages to members of both groups. This search ultimately yielded fifteen subjects over 90

days, ten men and five women. Six were freelancers for various outlets including specialized and popular scientific and environmental publications; the remaining nine worked in staff positions for news organizations ranging from specialized environmental outlets to small city and major urban newspapers as well as combination newspaper/online sites (Table 1). The subjects covered a geographic area from the U.S. northwest to the deep south to the northeast. Experience level ranged from approximately five years for a couple of younger reporters to a senior reporter with more than 35 years covering environmental news for the same news organization (although many news outlets had gone through multiple reorganizations over the time spans of some subjects' employment). Interviews ranged from 49 minutes to 92 minutes with a mean interview time of 1:03:39.

	Status	**M or F**	**Interview date**
1	Freelance, biomedical outlets; currently writing book	M	12/10/19
2	Staff, small city newspaper	F	12/18/19
3	Staff, public radio outlet	M	2/11/20
4	Staff, business publication	M	12/17/19
5	Freelance, multiple popular outlets	M	12/18/19
6	Freelance, popular and industry outlets	F	12/16/19
7	Staff, major city newspaper and website	M	12/13/19
8	Staff, major city newspaper and website	F	12/20/19
9	Staff, nonprofit state news magazine	M	12/30/19
10	Staff, public radio outlet	F	1/20/20
11	Freelance, popular outlets	F	1/23/20
12	Staff, nonprofit website	M	3/2/20
13	Staff, medium size city newspaper	M	2/20/20
14	Freelance, primarily industry outlets	M	2/11/20
15	Freelance, popular outlets	M	3/10/20

Table 1: Subject occupations and demographics

Locating a population of journalistic output for reconstruction

The standard method for reconstruction involves (Bruggemann, 2013; Godler and Reich, 2017; Reich, 2014; Reich and Barnoy, 2019) 1) assembling an inventory of all stories by each subject over a recent four week period, long enough to cover a wide range of material but recent enough for the subjects to remember the decision behind them. 2) using a random sample from this database as the subject of the reconstruction interviews. To follow this method, concurrently with the process of locating subjects, I assembled an inventory of all stories by each subject who consented to be interviewed (This was sometimes accomplished with the assistance of the subjects alone such as freelancers who usually posted recent stories on their personal websites, and sometimes augmented by database searches). Where possible, story collection focused on a recent four-week period (exact dates varied depending on each subject's output). This was long enough to cover a wide range of material but recent enough for the subjects to remember the decisions behind them. In a few cases, because of lack of recent material, it was necessary to locate older stories; however no stories were used as data unless subjects remembered them well enough to answer questions about them. I then took a random sample of up to five of these stories, located the measured facts (excluding uncontestable numbers such as dates or ages) and used these as the basis for the structured interviews. If a story was not about science or the environment or contained no facts that met the above criteria, I proceeded to the next story on the list, continuing until I reached a science or environmental story containing measured facts. Because so many routine statistics in the news such as economic figures (as well as such things as an analysis of experimental output in science news), may be reported without checking, based solely on the authority of the source, there was a risk that random sample alone might have yielded a high proportion of fact claims with an identical basis for justifying the decision to

include or not to check. To avoid this problem, I used a stratified sample that included at least one story in which the reporter had to verify a statistic, such as comparing a statistic with another statistical source (Barnoy and Reich, 2019). Stories in the former category were taken from a random sample while stories in the latter were based on suggestions drawn from the subjects themselves. Subjects were questioned about anywhere from a low of 1 to a high of 15 items, with fewer items for those whose interviews ran longer, in order to limit interviews to 60-75 minutes (Table 2; p. 46). This process continued until five stories had been selected and analyzed or until the story list was exhausted.

Semistructured and structured interviews

The next step was a semistructured qualitative interview with each of the first six subjects followed immediately by a structured qualitative interview about their stories. Each type of interview is a different kind of tool with a different function; both can serve complementary purposes for studies such as this dissertation, which approaches journalists' thinking both at the level of broad principle and of individual decision-making. Semistructured interviews address particular research questions through a conversation with subjects developed from a pre-existing protocol (Appendix A). A semistructured interview is simultaneously focused to elicit data on the questions at hand but also permit enough leeway to allow the interviewer to probe more deeply or probe for different answers when subjects' responses raise unexpected issues (Brinkmann, 2014; Lindlof and Taylor, 2017). This allows for both closed- and open-end questions such as "Why?" or "How?" and may venture into unforeseen issues when subject responses suggest it (Adams, 2015). Structured qualitative interviews, on the other hand, differ from both structured quantitative interviews (such as survey questionnaires) in which question

wording is identical for all subjects, and from unstructured interviews in which the subject's perspectives and concerns play the largest role and the interviewer's the smallest. Structured qualitative interviews combine some of the features of structured quantitative interviews such as standardized questions with some of the features of qualitative interviews such as flexibility in question presentation (Howe, 1988).

Questions focused on what Bruggemann (2013) identified as four components that characterize newsmaking at the microlevel: 1) the event itself that leads to the newsgathering process; 2) the trigger or indicator that makes the event manifest to someone in the editorial process, whether a journalist, editor or other actor; 3) evaluation, the process of deciding whether the event is newsworthy; 4) editorial context such as available resources, editorial policy and executive decisions, individual interests of particular actors. Based on these, all subjects were asked three questions about the statistics in their stories: 1) Did they remember the source of this statistic; 2) did they verify the statistic; 3) if so, how they verified it; if not, why they considered it reliable enough not to verify. These questions were always asked; however, wording sometimes varied in order to ensure that the intent of questions was always comprehensible for each statistic and the reporting situation in which each one arose. For example, when subjects explained that they used the same test of veracity for several statistics from the same or similar sources, it was not always necessary to ask the same verification questions for a sequence of related statistics in the same paragraph. Following Reich and Barnoy (2016; 2019), verification could include evidence of evidence such as a history of working with sources in the past. The two-part design was necessary for two reasons: (1) in order to gain subjects' confidence for what might be seen as an overly intrusive inquiry into their working methods; (2) to ascertain whether

subjects are verbally skilled and reflexive enough to address the questions required by the structured interview (Barnoy and Reich, 2019).

All interviews were audio recorded using Quicktime and transcribed. The semistructured portions of the transcripts were coded and the structured portions were systematically entered into an Excel sheet (See Appendix B for examples).

Analysis: semistructured interviews

Because the semistructured interviews covered much of the same ground as the interviews in Van Witsen (2018), the codes that emerged from that study were used as a form of first cycle coding for this analysis. Initial coding of that data began with simple topics with more abstract codes emerging inductively as data analysis progressed. 49 initial codes and 29 subcodes emerged inductively from the 535 separate coding units, with 194 analytical memos. Some codes that described very closely related concepts were eventually collapsed into a single code.

Coding of the semistructured interviews in this study began with the 2018 codes, but honed in on just one aspect of their professional lives, their reporting and verification decisions concerning the numbers in their stories. It began with these because the thinking of the new subjects was expected to have some continuities with my subjects from 2018, especially given the consistency of what I found in that study.

The semistructured interviews were analyzed at the sentence level and the word level (Chi, 1997; Simmerling and Janich, 2016) on the basis that neither single words nor whole sentences can express the entire range of meanings about the relationship among norms, values and practices. When results at two different levels of analysis contribute to the emergence of a single set of codes, this can function as a form of reliability (Chi, 1997). Because the codes that

emerged in Van Witsen (2018) were so consistent across subjects, this analysis began with these codes in order to search for congruencies and build from there. These include: (1) The professional nature of news work and the journalistic career; (2) Origins of numbers and their transparency; (3) Trust in numbers; (4) Numbers in stories: avoidable or unavoidable tension; (5) Statistics as culture: the role of context. (See results section for these codes in detail)

Analysis: structured interviews

For each subject, up to six stories (depending on the subject's output and story availability) were listed, with story title or headline on top and each measured fact in the story listed in the order in which it occurred in the story. Below each fact all available information from the structured interviews was listed including how the fact was discovered (where available) and what methods were used to verify it, including trust in sources. This method made it possible to study verification in depth for each fact and broadly across many facts in a single story, multiple stories by one subject or multiple stories by multiple subjects to search for patterns across the entire profession (Table 2). The mean number of stories analyzed per subject was 4.8. The mean number of statistics analyzed per story was 2.63.

Subject	# of stories analyzed	# of stories	# stats analyzed story 1	Story 2	Story 3	Story 4	Story 5	Story 6
1	4	6	6	7	5	1		
2	5	24	7	9	3	1	6	
3	5	8	1	3	2	3	2	
4	6	15	3	2	3	1	1	3
5	6	6	13	4	3	4	10	n/a*
6	5	28	3	4	4	1	4	
7	5	29	3	9	19	22	3	
8	5	13	7	8	3	3	15	
9	5	21	11	n/a*	6	3	10	
10	5	21	1	2	1	1	1	
11	4	27	1	1	5	1		
12	4	10	1	2	1	1		
13	5	22	2	3	2	3	1	
14	3	9	1	2	1			
15	5	17	1	2	1	1	3	

*Table 2: Verification results *story was compiled from secondary news sources and not primary source reporting*

Two units of analysis were employed for analyzing the structured interviews: (1) the entire story as each subject discussed the overall reporting task and (2) each numerical fact claim within the story. The two units were used in order to capture the dynamic relationship between an entire story and its components; whether, for example, the subject's detection of a single fact makes other facts in the story more salient, or whether confirmation of a single fact claim makes the truth of other fact claims more likely. Decisions about specific statistics in individual stories were compared to the codes that emerged inductively in the semistructured interviews in order to decontextualize data and derive new categories and concepts where applicable, through the abductive process (Richardson and Kramer, 2006). Unlike deductive reasoning, which operates by formal rules, or induction, which generalizes a likely explanation across multiple cases, abduction infers theory from best available explanation of existing patterns of rules that have

already emerged inductively. The goal was to try to link specific decisions about verification, trust and norms to the codes developed in the semistructured interviews in order to address the research questions.

The analysis was compared against the following operational definitions (Godler and Reich, 2017):

Verification and corroboration were operationalized as testing the claim with sources or comparing it with documents originating outside the source itself.

Trust was operationalized as relying on the accuracy or completeness of a fact claim independently of checking or corroboration.

The *source of each statistic* was operationalized as the individual or institution from whom or which the fact claim originated. A source can include every human, documentary, technological or organizational factor that contributed a meaningfully to the knowledge claim.

Source types were operationalized as each source's institutional, social, political, or professional role. Source types can be fit into such categories as officials or authorities, institutions and groups (including advocacy and interest groups), experts such as academics or authors and sources with no institutional role except to have witnessed or participated in an event (such as spectators, crowds, victims or exemplars). Data or numbers originating from a source are considered to be attributed to that source. A human spokesperson for a source is also attributed to that source, as is a document such as a video or a press release.

Norms, routines and policies were operationalized as informal rules based on division of journalistic labor, temporal imperatives such as deadlines or turnaround times, professional roles and precedents such as beats and specialties, institutional hierarchies such as the editorial chain of command, or professional values and expectations. For purposes of this study, they are present

when subjects perceive them as the basis for a decision. Individual journalistic agency operates when subjects perceive they could have acted otherwise. In accordance with Giddens (1986), the two are sometimes interlinked.

Analysis of these can clarify which statistics were trusted because of trust in their sources and what this trust was based on, such as familiarity or previous history with a source or the source's institutional authority or status. It could also provide a better understanding of when journalists decide they do not trust a statistic and what methods they use to cross-check.

RESULTS

Results below include 1) the codes developed through semistructured interviews and how these were developed; 2) examples of codes as they emerged in interview excerpts and discussion of these; 3) analysis of the forms of verification or its absence that emerged in the structured interviews including examples of these as they emerged in interview excerpts.

How new codes emerged from earlier codes

Because my previous research (Van Witsen, 2018) discovered robust categories that consistently explained the subjects' thinking about statistics, these were used as working categories for beginning the coding process. These were:

1. The professional nature of news work and the journalistic career: all subjects' conscious and continual awareness of the importance of the news production process in shaping the finished news product and of their role in it. These affected their views of moment-to-moment decisions on particular stories, how these related to other projects on which they were involved, and where these decisions fit with longer-term aspirations for themselves and their organization.
2. Origins of numbers and their transparency: even though all subjects recognized individual numbers could be problematic, the idea persisted that numbers provide direct access to a kind of truth not available from live sources or eyewitness descriptions.
3. Transparency and non-transparency: While subjects understood that individual numbers could be wrong or at least open to challenge, they appeared to grasp this on a case-by-case basis rather than theoretically.
4. Trust in numbers: The concept of statistical transparency can make it difficult to recognize the existence of alternative systems for conceptualizing categories and measuring them. Journalists

can sometimes develop a degree of skepticism about numbers, particularly on beat reporting that gave subjects sustained exposure to the details of number construction. However this was never assured.

5. Numbers in stories: avoidable or unavoidable tension? All interview subjects were aware of tension between the between the static abstractness of measurement vs. the dynamism, intentionality and tension of narrative. However they rarely took a firm side in the "numbers vs story" issue. This might be because of a recognition that the division is frequently not clear-cut (e.g. numbers sometimes possess narrative power of their own). Or it might be evidence of journalists' ideas about how they believed audiences respond to numbers.

6. The role of context: This usually meant the range of closely related references, frequently other numbers used by journalists to show connections between the numbers judged as newsworthy and other numbers from which they emerged. Supplying context frequently involves a search for additional facts, sometimes in the form of numbers, sometimes not, guided by the "gut instinct" on which journalists rely. Context sometimes included beliefs about the important status of statistics in general.

I began coding the first six interviews expecting to find subjects' thinking and decision making about verification would be influenced by the above codes and behavior patterns. New codes that were created were narrow due to the narrower focus of the new interviews, which were oriented mostly toward verification rather than general views of statistics alone. New codes that emerged were:

1. Career: following Tuchman (1972), Fishman (1980) and others who studied the news production process as a social system, all subjects were consciously aware of what motivated them to become journalists and the paths, direct and indirect, they followed to their present

positions. In all cases, the aspirations combined with circumstances to create their professional values and norms.

2. Sources: subjects' attitudes and beliefs about where to look for statistics as well as their actual behavior as they searched for forms of verification. These grew out of journalistic routine such as regular contact with a known group of sources but sometimes included nonroutine such as searches for new sources.

3. Journalistic routines: all subjects understood their role in the organized system by which news was discovered, reported on, verified and turned into finished editorial product. This was equally true for the freelancers and for staffers although the working routines of the two groups were of course different. Except for subject 1, who had partly withdrawn from the daily news production process, much of the subjects' time, imagination and energy were consumed by working to make this system function as well as possible. While they all were aware of its limits on time and resources along with the need to sometimes make expedient decisions, only subject 1, a freelance with long experience, had partially abandoned this system of news production to focus on spending long periods of time, sometimes as long as several years, with subjects in order to learn more about sources. All other subjects worked productively if not perfectly within this system. A few showed awareness of its limits when these were pointed out in interviews but had no interest in substantially changing their practices.

4. Journalistic routines and statistics: how the needs of the news production system affected choices about statistic use or verification.

5. Role of statistics and numbers: subjects' attitudes, ideas and beliefs about value quantified information contributes to the news.

6. Handling statistics and numbers: How subjects made decisions about verifying and using quantified information based on the above beliefs.

After the first six interviews were coded, the second six were likewise coded and codes and memos for the first group compared with those for the second. Possibly because of the narrow focus of the semistructured interviews, no new codes emerged in this second round of coding. Three additional interviews were then coded, for a total of fifteen. When no new codes emerged in these, saturation was determined to have been reached and interviewing was concluded.

Relationship between codes and journalistic practice

Career

Subjects entered science and environmental journalism by a variety of routes but almost always without planning or specific training aforethought. Many were intensely curious about scientific discoveries. Some had training in scientific research techniques but chose not to pursue a research career. Without exception, the environmental journalists were interested in environmentalism and the environment in some form. What united all subjects was some kind of previous, usually early, experience that gave them a close view of science or the environment. Subject 10 studied environmental science in college while subject 8 studied environmental sociology. However to call this a prerequisite for the role would be overstating the case. Subject 7's experience with environmental reporting predated any individual position including the present one.

Except for subject 1, nothing in any subject's preparation or attitudes prevented them from meeting the demands and needs of the news production system. This almost complete acceptance of their journalistic roles meant every subject's thinking and behavior was shaped by the

demands of that role as well as the demands of fact-finding. Most were aware of the tensions between the two but found them manageable. Subjects 10 and 12 had both spent time at environmental advocacy groups before turning to journalism and consciously understood the different expectations and standards for each.

> When I worked for (an advocacy group, the) rules were basically to break everybody else's rules. When I work as a journalist there are specific rules there about how you behave and how you present all sides and I think it's important to say all sides because there are very few environment stories that only have two sides. (Subject 12)

Subject 5 aspired to write exclusively about math for a popular audience but found it difficult to make a living from this narrow focus. Freelancing, for subject 5, led to longstanding, time-consuming relationships with ten to twelve news outlets while also spending some time offering stories to new outlets. A focus on audience meant constantly thinking about each story's potential to meet client needs, leading, for example, to more climate stories than necessary because clients wanted them. This meant mediating between a personal interest in scientific methodology and how science reaches its truths, vs a recognition that stories had to be sensitive to what audience did or didn't know about statistics: "…the disagreement that there might be two different methodologies means two different ways of assessing. That's pretty far in the weeds for an audience of 9 to 14 year olds…. They would just gloss over it and just keep reading."

Subject 13, with nearly three decades at the same outlet in a medium-size city, recognized how unusual it was to survive the massive changes currently being experienced by the profession. A stable work role growing out of a focus on the environment led to the creation of an extensive network of national and international as well as local contacts.

> I've had people tell me that I have access to more scientists or other officials who just don't talk to reporters. When I went to Greenland in 2008 … I was meeting this guy at Ohio State who's now he's one of the world's top Greenland researchers, he's now over in Denmark but he laughed when I met him. He said I ran your name up the flagpole with the Governor's office. He knew the governor at the time. And the woman who was his environmental advisor, he goes, you should be happy to know, you're one of only like two reporters in the state they've said that I should take my time other than the national media, Today show or 60 Minutes or things like that.

This extensive network functioned as a form of professional social capital that permitted the generation of many stories.

Sources

Strictly speaking, no journalist lacks access to news sources; in fact, the systems of information subsidy that form a part of journalistic routine provided all subjects an almost limitless number of science stories, at least in principle. Online services such as EurekAlert and Sciline offer access, sometimes on an embargoed basis, to a wide range of newly published research in all disciplines as well as scientifically credible sources to comment on it. Almost all subjects monitored scientific and other journals on their own, while many also subscribed to individual email news release lists. Subject 13, with long tenure on the environmental beat at a single outlet, received an estimated 400-500 emails a day even before counting letters, telephone messages and personal meetings. Some subjects also received private subsidies via professional relationships with knowledgeable individuals in scientific institutions who sometimes alerted them to relevant scientific research before it reached established channels.

The sheer scale and profusion of this medley of material, along with differing editorial priorities at different outlets, forced all subjects to create heuristics that could filter the full range of science news without completely ignoring lower priority news that had potential to grow in importance in the future. The actions of government agencies with an important impact on many lives regularly took priority, to such an extent that all subjects on the environment beat cultivated routine relationships with the agencies themselves and their key officials. Major decisions by these agencies were always considered important news.

Subject 2, consistent with the deference outlined above in *Journalistic Routines*, had high expectations of government and scientific sources: "I'd like to hope they're correct." Trust of government sources led to trust of their statistical output but since trust tended to increase as subjects gained more experience with their sources, this level of trust may not have depended on authority alone. The same was true for important national scientific institutions including government agencies such as the National Oceanic and Atmospheric Administration (NOAA), local scientific institutions such as universities or research institutes and scientifically oriented industries such as nuclear plants. Environmental NGOs of all kinds formed another important group of sources, but these were not treated as equally credible or having equal capacity to make news.

As in *Journalistic Routines*, trust was neither absolute nor consistent. Subject 13, along with some others, recognized that government agencies can have institutional agendas that affect the way they measure things. Subject 6 trusted science published in widely cited, high-impact journals and was more suspicious of open access journals, believing they lacked a credible review process. In a widely repeated pattern, this subject found a certain statistic through a news aggregation site, contacted the authors of the study, then looked for other researchers who were

familiar with the research but not involved in it. Subject 7, with long experience with environmental topics, cited some sources not simply through trust but because previous coverage of the same source had led to a deeper understanding of the controversies surrounding a topic including, occasionally, controversies over methods engendered by different perspectives on the problem (such as actual harm vs potential harm in calculating risk assessments). In addition, subjects' trust in some sources was not simply a matter of their previously established reliability but also derived from a source's monopoly position in the creation of certain kinds of measurements. When a measurement was only available through one government agency, subjects frequently felt they had no choice but to trust (although the same was not true of scientists with a similar monopoly on a scientific discovery).

Journalistic routines

With, again, the partial exception of subject 1, all subjects approached their work through routines; that is, their work exhibited a range of repeated behavior patterns for discovering news, verifying it and writing stories about it. The most significant difference was the expected one between staffers, with one set of editorial norms to satisfy, and freelancers, who frequently had many. Subject 6 spent a long time learning editors' preferences but also found the process of searching for stories was synergistic; that is, learning about a story suitable for one outlet sometimes turned up stories suitable for another.

> The different outlets cross pollinate each other. Like for instance, say I'm working on a story for [name of publication] and I really like putting together pitches for [name] because, as I said, there's a variety of journals that I like and my editor there likes stories from those journals. And so as I'm cruising through and scanning them or scanning, like,

> Newswise SciWire [an online distribution system for news releases about science] has like a weekly email with stories in it, Environmental News Network, has stories. … But as I'm scanning those, I'll see something and I go, "Ooh. [name of another publication] story, that would be good." Or I might do a story for [name] and say, "Hmm, that would actually, if I dig into it more, take a different angle that would be good for [name of third publication]." … And just the process of going through looking for story ideas for one outlet, will give you ideas for stories for the other one. It just kind of brings you up to speed on what's going on out there and environmental science beat.

Subject 5 spent a lot of time thinking about which stories would interest different news outlets based on knowledge of their different audiences. Subject 11 was sometimes given assignments by editors at different outlets but tried to exercise more editorial control by searching out stories based on a single scientific paper that raised interesting questions about theory or methods.

Subjects 7 and 13 both had long tenure at a single news outlet, allowing them to create an extensive network of national and international as well as local contacts. This form of professional social capital let them proactively generate many stories rather than simply reacting to the day's widely reported news.

> One of the things that we're planning on doing sometime next year is taking a deep dive into, plans to build two major, diversions on [name of river] that will be designed to move sediment from the river into wetland areas, open water to create new wetlands. and so that's something we want to do. It ends up, that numbers become a significant part of that process. How much sediment is the river carries down stream that can be captured? How much should that sediment actually get out, through the diversion to be available to create new land? How much sediment does it really require to build something. And, the

> fresh water that's delivered, along with that sediment, what effects are having on different fisheries and numbers involved in all that? (Subject 7)

Their ability to revisit important topics repeatedly gave them an institutional memory for previous facts that could be used as a basis for additional reporting. Subject 8, one of several environmental reporters for the same outlet, considered it an advantage that different journalists focused broadly but not exclusively on different areas of the environment, allowing different journalists to learn from each other. This much social capital was rare. Subject 12, working for an online environmental website with a small staff, was acutely aware of how these limits affected the ability to verify facts and locate new facts.

> Someone puts a year of their career in a published paper. And I take a day and a half to publish a story…. if a given Ph.D. has five research assistants working on a project for a year and I have three reporters who work on a new project every three days...I do not have the resources to even seriously think about that depth. That's why they call journalism the first draft of history.

Subject 1, unlike all other subjects, had made a conscious decision to withdraw from daily journalism to focus on longform biomedical stories. Standing outside existing newsroom routines made it possible to spend more than a decade following a single topic and build close relationships with the human sources before deciding to report it as a story. This made it possible to avoid overcovered stories and to learn science in enough depth to have developed some independent judgment about the science. One example: the belief that claims about the effectiveness of HIV vaccines were exaggerated, but “I’d be crucified if I tried to write about this.” It was unclear whether this depth of knowledge enabled greater perspective or cost some measure of independence from sources. Other subjects discovered news and verified it through a

narrower range of practices that allowed them to satisfy the profession and the organization's expectations, including productivity and verification, but this had important implications for the meaning of verification with regard to statistics.

Journalistic routines and statistics

The subjects of this study almost always relied on official authority or expertise to determine what counted as usable quantified knowledge. Although none offered any official newsroom rules for the practice, this reliance remained a tacit norm by default, almost an axiom for any research about public events such as existing environmental problems rather than new scientific knowledge. Subject 9 considered government data a credible first place to look when beginning research. Subject 10's routine practice of starting the search or numbers with official or government agencies, grew partly out of the institutional authority of their official position and all that entailed.

> The story I just did I used, a lot of information put out by the Great Lakes Fisheries Commission. They do a lot of studying and working on, and it's their job to control sea lampreys in the Great Lakes. So they have a lot of data on dams, barriers and like the importance of those in the region. I would say government science agencies are like my number one go to. I like to cover climate change. So like, NOAA is my number one, cause they keep all of that temperature and precipitation data for years and years and years back.

However axiomatic, the practice was not arbitrary. Like subject 10, above, many subjects relied on government agencies (such as the Census or the weather reports provided by NOAA) because these agencies' official rationale was to make certain kinds of measurements, giving them a *de facto* monopoly on some statistics. Subject 6 frequently cited official sources because of their

existing reputation for credibility, calling this a necessary shortcut for freelance work. Even the staffers, such as subject 2, relied on official numbers in the absence of alternatives: “they are the officials on it….I just feel we’ve got to go with what they have.”

Journalistic routine did not mean unthinking acceptance of official measurements or official definitions of what counted as a fact. Nor did it mean the subjects were open to anything. All subjects had access to sources that extended beyond officialdom which they used both to develop stories and to verify official statistics. Subject 4 worked at a business publication that had developed special resources to investigate the conditions behind production of business and accounting numbers. This allowed for reporting on science and high-tech firms in unusual business depth while still not affording the same level of resources to investigate purely scientific statistics. Subject 11, like subject 5, was interested in problems of scientific methodology and was sometimes able to dig deeply into scientific data as a purely intellectual challenge. Subject 15, working for an outlet that covered science for a popular audience, felt freer than others to examine the methodology behind numbers.

> There were no, let's say, sources that you just knew as entirely unbiased. Everybody, whatever the source, there's some kind of assumption that goes into those statistics or that research or that model or whatever.
>
> Q:
>
> How often did you look into the methods behind the number creation?
>
> A:
>
> I mean all times. … I mean pretty much always. You would always look at the methods for (name of publication) I should say. Not true at other places necessarily.

While not all numbers were verified, when the subjects did verify, this usually meant searching for alternative sources doing parallel or similar work to see if the numbers generated by the two were similar in magnitude or range. Those with broader source networks did this more easily; however, most subjects worked to extend their sources when verification called for it. Subject 2 developed an interest in "blue carbon," the term for carbon stored in coastal and marine ecosystems, and over time, was able to develop new local and national sources of information that allowed for cross-checking statistics.

Unless the subjects' sources disagreed on ideas, perspectives or methods in ways that became manifest to them, they almost never questioned the conceptual basis for number creation. Subjects 5, well versed in scientific methodology, reported only on methodological disputes that were manifest rather than latent. Subject 11, also knowledgeable about science, occasionally found ways to work outside journalistic routines and had published stories critical of the methods used in particular research.

> A huge part of what I try and do is explain how the decisions were made to collect the numbers in a certain way. What question is being asked in the first place? The decisions that were made in quantifying something in a certain way, what was left out of those decisions? All of those are essential to understanding what statistics mean. …
> I'd be looking at the data… if it's open and take a look at how the analysis was done or I'll read the statistical methods in depth. Because of the work that I do in the replication crisis, I'm inherently quite suspicious about how statistics have been used in the paper and how strong the conclusion can be based on the statistical methods. I'll be looking for things like whether there are analyses that work reportage that they conducted within, they didn't report the low results from them. I'll be looking at whether their analyses were

> preregistered or whether they kind of post-hoc made decisions to like drop participants or otherwise remove outliers and that kind of thing. Whether they had a small sample size within the analyzed six different, depending.

Subject 11's activity represented an exception. The general pattern was that however much statistics may have been created things (Boyd and Crawford, 2012; Gitelman, 2013; Hacking, 1990) the basis for creating them was never considered news by itself.

Role of statistics and numbers

All subjects, without exception, believed statistics play a fundamentally important role not only in reporting the news but in the definition of what constitutes news in the first place. Subject 9 talked about the importance of quantification as a way of establishing certain kinds of meaning such as progress toward a goal.

> Whether it's temperature or a trend in polluted hotspots, it's really the only way to gauge progress or lack of lack of progress on, really, any issue at all, is to have credible data and credible numbers to say otherwise. Otherwise you're stuck with the only anecdotes and you're not telling the reader the whole story to the best of your ability.

Scale, in the words of subject 11, makes a unique contribution to understanding the meaning and significance of something. This is closely related to what counts as news, especially in the sciences, which almost always involve measurement. Discussing a story about the decline of critically endangered white-backed vultures in Africa, subject 11 used numbers to "sketch out for readers how big the threat is, trying to quantify in a way that they can understand if this carries on, what is likely to happen to vulture populations." Writing stories about the replication crisis

had made subject 11 unusually attentive to data handling and questions of how strongly a scientific investigation's conclusion was based on its data.

Subject 9, like some others, considered anecdotes a suspect form of knowledge claim. Systematic knowledge was considered the way to go. This was particularly true for subject 4, a financial reporter for a business publication who took for granted the existence of an audience of executive decision makers familiar with and sympathetic to quantitative thinking. Frequently this was related to ideas about the scope and magnitude of an event, particularly profits and losses. Subject 2 said reporting the scope and magnitude of something is an important way "for painting the picture for people. I'm giving them something they can actually think about or maybe relate to."

These beliefs were modified with two caveats. The first was that news production is always a limited, routinized process. Simple concepts of the significance of numbers could not always be avoided, particularly when the reporting was routine (such as monthly economic figures) or amid the uncertainties of daily news production. Subject 7 had been querying the Coast Guard for figures on an oil spill:

> In most cases, I'm sort of forced to trust them cause they're the only ones who are capturing that information. It depends…. in this case, it was much easier because he was on grass, so that, they were able to capture it fairly well.… there's another facility offshore, I can't remember the name of it, that's been leaking since 2004 every day. And, there are serious questions about the reliability of estimates that both the industry and the Coast Guard and anybody else's making about that, cause they just don't really know…. Most times like this, I started with a press release and then called the Coast Guard to ask more information about it. And if I'm lucky, they get back to me.

Q:

You're the major newspaper in [name of city]. How are they not going to get back to you?

A:

Because they have other things that they're doing or it's too late in the day. I don't know.

In this case the limits on a source (too busy, otherwise occupied or a simple error) combined with journalistic routine (the need to complete a story by deadline) to keep a potentially important statistic out of reach. If the story had been important enough to have merited continuing coverage over several days, this journalist might have persisted with the Coast Guard and eventually located the missing data. If it had not, and the subject moved on to other stories, this fact might never have gotten into print. This is an example of what Reich and Barnoy (2016) called "the simplicity, linearity, and structured nature of regular news reporting processes as well as their idiosyncratic, chaotic, and surprising elements, which cannot always be simply and smoothly reduced into rubrics and categories." (p. 3)

The second caveat was that many subjects knew numbers alone did not always speak for themselves. Subject 5 felt the meaning of what numbers were measuring had to be interpreted in order to be useful. Unusually for this group, subject 5 sometimes used climate stories to investigate how climate models were created and how these yielded certain conclusions. Subject 10 did not believe numbers were more or less objective than other forms of knowledge claim such as quotations but felt they allowed a story to go beyond anecdote, as long as they are used responsibly.

When you're writing a story, one or two people can come to you and say, hey, this phenomenon is happening. And I think there is a lot of reporting or has been a lot of

> reporting done with that method, and that's great. But I think it's even stronger when you can say, hey, I've got one or two people who are experiencing this phenomenon. And I also have all these numbers that show that this is like a trend happening, all over the county, all of the state.

This is an example of the idea of context, the relationship of a new event to previous events or a previous class of events (such as weather) to establish a pattern.

Handling statistics and numbers

These multilayered concepts of statistics did not arise in a vacuum. Subjects had to make critical, sometimes quick decisions about them and incorporate them into news stories in accordance with the norms of the media institutions for which they worked. Frequently their decisions had to meet multiple criteria including when to rely on some process other than formal verification in deciding when to trust numbers and use them. The subjects' handling of statistics varied over a limited range of practices, which were always tacitly understood and never subject to explicit formal policies. Subject 6, a freelance, knew that different news outlets have different demands but could not say what these were because none had issued any standards for acceptable handling of statistics. Subject 7, the staff environmental reporter with lengthy tenure, also was unable to cite any formal norms for statistics handling but believed numbers almost always emerged during the reporting process, that is, the process of discovering and verifying the full range of facts that takes place after the initial determination that an event is newsworthy. Subject 8 backed the assessment that statistics emerge during the reporting process but also thought data sets were richer and more interesting than simply the point they are used to support in a single story and could sometimes lead to other stories.

A few subjects occasionally wrote stories explicitly focused on the problems with a newsworthy or controversial statistic. This required them to cite the questionable statistic not for its ability to illuminate the news but to call attention to its problematic status. Subject 11 wrote a story about credibility problems with the World Nuclear Industry Status Report. This story included statistics from the Report saying new renewable power is capable of reducing more carbon emissions than nuclear energy, which were included in order to provide context for the critique. This might be called a rhetorical use of statistics. Subject 11 was one of three subjects who occasionally reported on the internal issues with scientific and statistical methodology, a rare enough event that it cannot easily be explained by reference to journalistic routines.

Reconstruction, verification, and statistics

While the semistructured interviews yielded a clear picture of subjects' principles about verification, they do not show whether their actual working practices reflected their stated beliefs. Verification research recognizes that not all facts in a news story are verified through some formal method such as cross checking and that there are no overt tests for which facts merit the verification process and which do not. Because these processes frequently appeared "mysterious" and "arbitrary" to past researchers, Godler and Reich (2015, p. 2) proposed a practice-centered perspective which seeks to show how journalists' ideas of truth and verification grow out of the logic of their ways of working such as different source characteristics and journalists' relationships with them. The reconstruction method works in granular detail on a case-by-case basis to show how verification principles are applied (or not applied) in the chain of decisions that lead to actual verification practices when a story is created. A particular advantage of reconstruction is its ability to show what happens to journalists' principles amid the surprises,

contingencies and pressures, the actual flux, of daily news production where events emerge suddenly, change quickly and unexpectedly, and verifiable truths may be difficult to discriminate from rumors, propaganda and confusion and background noise (See, for example, the interview with subject 7 on p. 62). The findings described here represent the first effort to answer these questions in detail with regard to statistics.

In almost all cases the subjects were able to recall their work processes well enough to allow them to be "reverse engineered" (Reich and Barnoy, 2016 p. 2). Because the structured interviews questioned subjects only about the quantified facts in their stories and disregarded other facts, the results do not represent a complete picture of a story's development. In addition, some stories contained only a limited number of statistics, resulting in too little data overall to allow for a quantitative analysis. However all the subjects recalled the stories under investigation well enough to enable them to answer the structured questions in all but a couple of cases. The richness and detail of subjects' description of their thinking and decision-making showed patterns that were broadly consistent with the methods and principles described in the semistructured interviews. Several subjects even described their decision-making in enough detail to recognize occasions when their actions fell short of their principles. The robustness of these patterns emerged at three levels: in multiple verification decisions within a single story, across different stories by the same subject, and frequently across different stories by different subjects. The patterns included:

The essential role played by official and authoritative sources. As might be expected from the "sources" code, this pattern existed at all times, across all subjects and in all stories. There were no exceptions. Without always stating this in words, all subjects took it for granted that quantified knowledge claims were found solely among sources who were recognized as formally

qualified to produce them, whether through their official position (such as a government agency tasked with creating particular statistics) or their knowledge, qualifications and training (such as scientists or academic experts). Subject 2 explained the reliance on the National Weather Service along with a regional agency that monitored snowpack level by saying, “Their official capacity is to be the experts on these topics.”

While citizens and other non-elites may be recognized as legitimate sources of other kinds of knowledge claims (e.g. voters, victims, or participants in events) the subjects granted statistical source status to elites exclusively and then only to elites with appropriate disciplinary qualifications; i.e. atmospheric scientists were not solicited for knowledge of cancer statistics.

Limits on the role of official sources. Despite what Lugo-Ocando and Brandão (2016) found, all subjects understood that statistics could be checked and verified; that is, none of the interviewees believed quantified knowledge claims could not be argued with. Formal verification, as operationalized on p 40, operated along a single dimension: it meant further reporting to find other sources of similar statistics that could be compared with the first statistics to determine whether the different statistics aligned and if so, how closely. Subject 2 learned about owl-vehicle collisions from a wildlife center and verified some of the numbers and locations of these collisions by comparing them with figures from a second wildlife center. This same subject verified figures on the whale population from an environmental group in the northwest by comparing them with figures from another group also dedicated to whale conservation, then compared them again with figures from the National Oceanic and Atmospheric Administration (NOAA). The respondent said: “I was gathering a lot of data from all of them. And then I felt comfortable enough that multiple sources were in agreement.”

Comfort level is subjective of course. When figures from different sources were of the same order of magnitude, they were considered verified, although this did not explain why subjects published one of the figures rather than another. Subject 4 gathered figures on the size of the commercial space industry from an investment banking firm, then compared them with similar estimates from three other firms. The original estimate was cited when it turned out to be about the median figure for the entire group. In addition, subject 4 considered the first firm's research "pretty thorough" but did not explain why the original figure was cited rather than any of the other three.

Subject 7 searched for other experts studying the same general environmental phenomenon, then compared their numbers but was sometimes able to skip this step due to long familiarity with multiple sources studying the same phenomena. Most subjects understood how explicit agendas could affect the counting process but were more likely to use this caveat for statistics that originated with advocacy or interest groups, including environmental groups. Only a minority, such as subjects 1 and 7 recognized that official sources such as government agencies could have agendas of their own that might affect their statistical output.

Formal verification, for statistics, therefore, meant two things: 1) searching for experts familiar with the number in question but not directly involved in its production who could comment on either the statistic or the validity of the research that produced it; 2) searching for alternative sources of the same number (again from qualified but uninvolved experts), then comparing the two. If the numbers matched or were within the same range, the first statistic was considered accurate enough to use in a story. Almost all subjects performed these verification operations at times, and some performed them frequently. Journalists with a wide range of sources or long experience with reporting on a recurring statistic (such as pollution levels)

sometimes became familiar with the methods behind numbers production, a form of professional capital that allowed for easier or more frequent verification or allowed them to recognize anomalies or unexpected changes in the pattern.

When formal verification was not performed

Subjects were less likely to verify statistics originating with an official or authoritative source if they had worked with the source in the past and found it to be reliable. There were multiple examples of this across the entire range of subjects. Subject 15 cited the Keeling curve measurement of CO_2 concentration in the atmosphere without further verification, based on knowledge that this statistic had an unbroken record going back to 1958. Subject 2 once again reported feeling "comfortable" citing a figure on annual release of Chinook salmon from the state Department of Fish and Wildlife based on a history of working with this agency in the past. Since this procedure is not based on source authority alone but on the source's past performance, it can be considered a practice more subject to evaluation of some kind.

When subjects judged they had sufficient history with a source to justify citing a statistic without going through formal verification, they sometimes applied it on a blanket basis to several statistics in the same story originating with the same source. Subject 9 cited estimates of the effect of tariffs on soybeans based on a past history of working with the U.S. Farm Bureau, then cited a similar figure about corn for the same reason without developing a separate rationale. Subject 7 cited several figures on local toxic chemical emissions drawn from a U.S. EPA database despite recognizing that these figures were often "either reviews by the state or federal government of information provided by the industry or reporting by the federal or state government of information that they had collected."

Statistics originating with official sources were also sometimes not verified when a source such as a government agency was believed to have a monopoly on the measurement in question. In these instances, subjects explained that the monopoly status of the number made it uncheckable and simply attributed the statistic to the source. This presumably allowed audiences to make their own judgment about its credibility (Whether audiences actually did so and how they accomplished it is unknown). In addition, some subjects such as subject 12 defaulted to monopoly sources when they believed checking wasn't possible for short turnaround stories produced under extreme time pressure. This happened even though subject 12 recognized statistics from government agencies such as the National Park Service could have institutional biases of their own. The process of defaulting to monopoly sources appeared more likely when the statistic wasn't considered central to the story's news value. e.g.:

> Q:
>
> How did you decide to trust the figure of a population of 7,000 for [name of city]?
>
> A:
>
> Trust that the US Census is the best estimates, tends to be the really only authoritative kind of people. (Subject 9)

Subject 9 twice cited population figures from the US Census without checking, despite recognizing that Census counts have their own politics.

All subjects used these short cuts at least occasionally, including those who understood the principle that a single statistic didn't represent unchallengeable knowledge and could be verified or falsified. Subject 13 relied on comments by an official of the U.S. EPA that radiation levels at a proposed site for storing radioactive wastes were at or below background levels. This subject recognized that, with more time, better data could have been located through the Freedom of

Information Act but nevertheless made a judgment that relying on a single quote from a single official did not represent a serious risk. Subject 14 cited a figure on the cost of hydrogen fuel based on general experience reporting for chemical industry publications: “so I didn’t have a lot of reason to question that.”

Subjects who repeatedly covered the same statistics over many stories occasionally learned enough about the methods behind them to factor these into their news judgments. Subject 7, with decades covering the environment at the same outlet, published stories about a controversy over different methods of risk assessments by the Army Corps of Engineers. At times this subject made a decision not to include challenges to the Corps’s methods, believing they weren’t accurate. Reporting with this much conceptual depth was rare. In general, subjects did not inquire into methodology behind a statistic unless they perceived a conflict in methods between two or more producers of the same statistics, which made the judgments and values behind the methodology visible. Elsewhere, subjects simply noted the differences where and when they became apparent and included them in their reporting when they considered it appropriate.

This range of practices was never perfectly consistent and was frequently subject to the routines of news production including compressed time frames, limited budgets and the challenge of comprehending events that changed at the same short time scale as the processes of trying to report on them. When questioned, they understood the limits of simple reporting conventions such as defaulting to official statistics but never hesitated to use them. Subject 15, who used the Keeling curve measurement of atmospheric CO_2, also, in a different story, reported on the relative cost of stormwater vs fresh water by citing a figure from a government agency, the Los Angeles Department of Water and Power. Recognizing that this agency had interests of

its own that might influence its counting methods, he cited the figure and attributed it to the Department, reasoning that audiences could decide its credibility for themselves.

DISCUSSION

The results described above provide a robust basis for addressing this study's research questions. Consistent with Van Witsen (2018), belief in the special epistemic status of numbers existed but was limited to particular circumstances. Subjects verified numbers in many cases but trusted in many other cases. Pure trust in numbers operated only in particular, circumscribed ways and was sometimes difficult to distinguish from trust in the sources of those numbers (See CONCLUSIONS for a theoretical explanation). Norms and routines of news production, always applied tacitly rather than overtly, did not dictate a single definition of what constitutes an accurate statistic or even a single test. The conceptual issues behind statistics creation studied by such researchers as Lugo Ocando and Brandão (2016) and Lugo-Ocando and Lawson (2017) were rarely raised in individual stories even when interview questions showed that subjects understood them. They were raised only when an external controversy made them newsworthy; subjects never initiated these challenges themselves. Concepts, in other words, may have shaped statistical output but it was the output that made news, not the concepts. This is consistent with previous findings (Van Witsen, 2018) which found journalists' understanding of the normative, negotiated status of category definitions and the imperfections of the counting process took place case-by-case and had no recognizable basis in principle.

The first research question inquired into which statistics journalists considered newsworthy enough to include or not include in their stories. All subjects interviewed believed statistics were a valuable component of reporting, essential to establishing certain kinds of meaning that is probably inseparable from general concepts of newsworthiness such as a change in the status or phase of something (Fishman, 1980) In journalism, such changes are frequently established by measurement (e.g. gross domestic product or an election outcome or level of CO_2 in the

atmosphere) although not always (e.g. a trial verdict or the death of a prominent person). Because change in science, including environmental science, usually involves measurement, the subjects of this study found numbers a natural ally for their kind of reporting. In the initial stage of story detection, newsworthy statistics were those that alerted subjects to the existence of any of these changes. In later stages of the reporting process, newsworthy statistics were those that added either detail or additional meaning to the original change or those that complicated or countered it. Non-newsworthy statistics were those that tended to confirm existing knowledge about something rather than change it.

The second research question inquired into which types of statistics journalists believed need checking and which do not. All subjects regarded statistics that were central to a story's import as contestable as long as they perceived competing statistics to be available. However, statistics only compiled by and available from a single authoritative source (what this study refers to as "monopoly statistics") were considered uncheckable and were frequently cited on the basis that they could be attributed to these sources, with audiences presumably left to decide on the value of these numbers themselves. Statistics originating with scientists, businesses or NGOs were treated as falsifiable if alternative statistics could be found to compare them with. Statistics that played an important role in a story's central significance were more likely to be checked than supporting statistics or statistics that appeared to be unchallengeable (Subject 9 stated, "If it's something basic about how many inland lakes are in the state of Michigan, you can probably trust it. That's still about the same. But if it's a case of how many polluted sites the state is responsible for cleaning up that had been abandoned, that number is probably going to be changing even by the week or month."). All these criteria were flexible in two ways: first, when subjects were working with fewer resources or time, routine or monopoly statistics were more

likely to be cited solely with attribution to their sources but without independent checking. Second, when subjects had found a source sufficiently reliable in the past, they cited new statistics from the same source with no more verification than an attribution. This usually happened only after the subjects had a history of previously working with the sources although how much history was required varied.

Research question three inquired into how the status of a statistical source affected judgments about its credibility. While some subjects said they examined the credibility of a statistic itself, most subjects looked to sources first and tended to grant more credibility to government sources, especially when they were monopoly sources. Subjects with experience covering a particular topic were more likely to try to verify statistics including some from official sources; in other words, the special status granted to government statistics was not absolute but subject to additional tests.

Research question four inquired into the role of perceived organizational norms on verification of statistics. None of the subjects of this study could cite any formal rules for verification of statistics by the outlets they worked for, nor of any rules for verification at all. This did not, however, mean the subjects thought they were free to do anything they wanted. Informal, tacit beliefs and understandings about verification were frequent and subjects made decisions in full awareness of them. In spite of variations from outlet to outlet, these were governed by broadly consistent conventions (discussed further in RQ 5).

When subjects described the process of working with editors to verify statistic, the standards they described closely approximated those of the subjects themselves. Disputes over accuracy of a statistic meant disputes about how to make existing verification standards work as well as possible rather than disputes over the standards themselves. For example while subject 7

sometimes debated editors over how many numbers in a story were too many, neither disputed the importance of numbers in the news.

Research question five asked what formal tests journalists used to check the accuracy of statistics. No statistic could be cited without reasons, but not all reasons were formal tests of verification. Two formal tests were found. The first was to search for alternative sources measuring or counting approximately the same phenomenon, usually by the same methods. If these agreed or were close, the statistic was considered verified. The second, closely related to the first, was to locate and interview someone familiar with the production of the original statistic or doing roughly similar scientific work. If the second source considered the statistics produced by the first source accurate or credible, it was considered verified. Reliance on evidence of evidence was also frequently used such as previous history with a source. If a subject's previous use of statistics from this source had never been challenged, that counted as justification.

The overall picture that emerges from this study is of subjects who had varying degrees of professional social capital to help them verify information but could almost always use the resources available to them to perform levels of verification they judged they needed. This varied considerably for journalists working for outlets with different needs, such as a trade journal or a major city newspaper. All subjects believed measured information performed several important roles in the news not possible with other kinds of information, particularly as ways to judge the scale of an event or progress toward a goal. Verification of statistics worked through tacit but powerful rules that still gave all the subjects many choices about what counted as verification and how to perform it. The most important of these was a heavy reliance on official, authoritative or expert sources of statistic. No subject accepted the word of these sources as gospel. Subjects

were always ready to check the accuracy of quantified knowledge claims but this was not always possible or feasible. When subjects could not verify official statistics they turned to their previous history with sources. If they judged statistical sources reliable in the past, they used this as a basis for trusting new information from the same sources without checking. These verification tools were not used at all times; in at least some cases, statistics, especially routine or regular numbers issues by authoritative sources were published with an attribution only. The same was true of sources believed to have a monopoly on some figures. The frequent contingency and unplanned elements of daily news production meant none of these findings held true in all cases.

CONCLUSIONS

This study of statistics in the news was the third in a series examining how journalists think about and use statistics in their reporting and writing. It grew out of the concern that many journalists believe measured knowledge, as a concept, cannot be argued with, ruling out alternate ways of thinking about an issue. These essential but imperfect numbers shape public perceptions of the size and scope of the things they were created to measure, including the social problems much news focuses on. But the simplification of journalism means the calculations and judgments behind the numbers take place offstage so to speak, hiding their contingency and the choices made in creating them. Properly applied, the findings from this study can help journalists acquire what might be called "data literacy," an understanding of the ways seemingly neutral and objective statistics are constructed and what can be legitimately inferred from them. This kind of insight can especially help people and perspectives that may be marginalized in the news learn how to challenge and question official views of reality.

This study used mixed methods to examine how science and environmental journalists' principles and beliefs about statistics relate to their decisions to verify individual statistics within individual stories. Behind these issues are larger questions about what the media routinely treat as uncontroversial and therefore not worth checking. Lugo-Ocando and Brandão (2016) and Lugo-Ocando and Lawson (2017) went so far as to argue that journalists are oblivious to the constructed nature of many statistics and habitually defer to official authority's definition of what counts as good measurement. This causes them to unquestioningly pass along arbitrary but deeply embedded ideas and perspectives as raw facts.

The findings from this study do not disconfirm that argument but they do complicate it. The subjects of this study seldom passed statistics along to their audiences without some form of

verification. Frequently they tested authoritative number claims (along with many that were less authoritative) by searching for alternative number claims to compare them with. When alternatives were not available, they did not automatically defer to authority but made individual decisions whether to trust or not, based on that authority's previous record and their own (or their colleagues') previous history with it. Subjects who had spent substantial time with a particular statistic sometimes understood the conceptual and methodological issues behind statistics creation even if they couldn't always apply them. Three subjects with a particular interest in scientific methodology (too many to be considered mere noise) sometimes examined statistical methods including how data were defined and gathered, which cases were dropped or retained, and justifications behind the analysis. Subjects with long experience covering a particular story occasionally became close enough to the statistics production process to understand the conceptual basis for different risk models.

In short, the subjects of this study sometimes recognized the basis for data creation, sometimes checked statistics through comparative methods that ignored conceptual issues, and sometimes passively accepted statistics on the say-so of authorities, just as Lugo-Ocando and Lugo-Ocando and Lawson charged. Far from not recognizing (or apparently caring about) the contingencies of statistics, these subjects engaged in an amalgam of practices that cannot be "simply and smoothly reduced into rubrics and categories." (Reich and Barnoy, 2016 p 2)

Although the behavior of these subjects upholds researchers (Fishman, 1980; Gans, 2004; Tuchman, 1973) who argue that journalists depend on norms to do their work, their norms encompassed a broader range of processes than those chronicled by these researchers. Subjects were continuously aware of institutional processes such as the editorial chain of command and professional values, but rarely perceived these to be a threat to their ability to act in ways they

thought appropriate. This does not, of course, end the debate; social control may still exist even if the controlees do not perceive it as such. It might, however, raise questions about what the debate over the power of journalistic norms and routines is about and what is at stake.

While these findings show the continued importance of trust and authority in determining what journalists treat as true, they also show the importance of Godler and Reich's concept (2015) of evidence of evidence. Neither of those concepts overlaid the subjects' practices on any one-to-one basis. The theoretical implications of this spectrum of behaviors will be discussed in the next section.

Theoretical implications

The theoretical implications of this work are built around Reich and Barnoy's observations (2016) concerning certain axiomatic premises about journalism. One of these is that processes always matter, meaning the truth and value of journalistic knowledge production can't be separated from the way it's created. The second is that these processes are simple in principle but always complex in practice. Many of the components that make up news stories as well as the processes used to discover them are concealed in the final product, "under the hood," so to speak. These can include undisclosed and unmentioned anonymous sources who might have served as second- or third-level forms of verification for the visible facts, or private information subsidies such as leaks and tips that make it possible to detect the existence of the news event in the first place, or later, unexpected developments or discoveries that force journalists to revise their thinking about the reliability of existing sources or their significance, or the relationship between previously discovered facts. Reich and Barnoy called these processes "complex and simple at the same time," (p. 3)

For my theoretical purposes, two issues emerge from this. The first is that even though the reporting process theoretically has distinct stages such as fact discovery, additional fact gathering, fact verification and creation of the finished editorial product, the actual conditions under which news is produced rarely make it possible to easily distinguish these stages from each other. Sometimes two or more stages combine into one. Sometimes the process is recursive, such as when information learned at a later stage of the reporting process forces a rethinking of some of the conclusions or even the importance of what was learned in an earlier stage. In addition, journalists never know when they can work in a straightforward manner and when they will have to deal with one-off events such as interesting but unverifiable information or surprises such as a sudden acceleration of their turnaround time for a story. The broader picture that emerges is of people formally tasked with finding and verifying facts who have to accommodate themselves to never knowing whether they can verify anything to the standards they aspire to.

The corollary issue is that if even the most routine reporting assignments might at any time require their creators to deal with both routine and nonroutine events and facts, flexibility of epistemic standards is almost a requisite for the job. I encountered subjects who did detailed investigations of the conditions under which a scientific finding was produced as well as subjects who routinely cited scientific authority because, as they put it, they could never do in one day what a PhD with a research team did in three years. In almost all cases, my subjects knew they could default to such a simple epistemic standard, even as they recognize other standards are possible.

Based on this I argue that journalists must always be prepared to continually evaluate not only each story but each fact claim within it during the reporting process, and adjust their verification standards to the shifting demands of each one, to the relationship between them, and (possibly

with editors assisting) to the relationship between one story and all the others in the day's news output. Experienced journalists may be so practiced in this kind of epistemic flexibility that they are able to make these adjustments intuitively, without conscious thought. The complexity and simplicity lie in the fact that in all cases, journalists know they can default to the simplest of expected routines such as attributing fact claims to authorities, even when they recognize more detailed reporting might yield a more precise picture of the measured facts (See, for example subject 13, p. 71).

The significance of this lies in the fact that, as one of the most important actors in the creation of the news product, journalists are central actors in negotiating both of what will be considered true and what standards will be used to verify those truths. Neither of these remain stable for very long. They can include anything from defaulting to authority, to defaulting to monopoly, to evidence of evidence, to careful investigation of statistical methodology. Caught at the center of these shifts, journalists must continually compare what they know about a truth claim, what else they believe (or suspect) might be known, and the resources available for verification. This makes their task both easy and difficult at the same time because their own individual freedom to decide what counts as true is never absolute, only a component in a larger process.

I propose that what the story reports as measured truth grows out of this interaction, in a way that should be testable. In their verification research, Barnoy and Reich (2019) used reconstruction to measure whether different types of news sources were more or less likely to be checked. A similar method could also be used to assemble a population of journalists, then ask them about a sample of measured and non-measured knowledge claims in their stories. Questions about each story-item could include about whether they used verification at all, which method of verification they used (cross checking or evidence of evidence), time frame for story production,

availability of sources, source type (government, expert, academic, NGO, advocacy or interest group, PR, etc) and different levels of perceived significance of the fact within the entire story. If correlations emerge between, for example, source type and form of verification or time frame and verification, this could be used to develop a theory of how epistemic standards for statistics vary depending on the presence or absence of other variables. It would also be possible to see if epistemic standards were different for statistical than for non-statistical facts.

If predictable relationships emerge between these variables, these might help explain the seeming contradiction between two discoveries from this investigation: first, that the subjects all regarded statistics and quantification generally, as having a special epistemic value, particularly for their tasks as reporters. Second, that the special status of statistics did not automatically exempt them from the professional habit of journalistic skepticism; in fact subjects were frequently prepared to test their accuracy, at least according to the verification standards they could perceive. To that extent, statistics were "sacred" and "profane" at the same time, in ways that matched Reich and Barnoy's observations (2016) about the continually shifting standards of the reporting process. Nothing in the finding that journalists frequently try to verify statistics contradicts previous findings about the continuing influence of cultural beliefs about numbers. Conversely, journalism's belief in statistics and confidence in their meaning does not completely explain where the particular authority of official statistics lies when journalists defer to it. Are they swayed by the official status of official sources or by the cultural status of quantification and its symbiotic relationship with journalism? It is not clear that these are separate phenomena.

Limits and suggestions for further research

Limits

This study leaves several questions unanswered. First, at a time of great changes in traditional journalistic norms, including an erosion in the monopoly professional journalists once had over newsgathering and a greater role for interactivity and for citizens as journalists and as sources (Deuze, 2008), it does not explain why these changes did not apply to reporting about statistics, which drew entirely on traditional sources. While statistics were not always trusted and no single statistic, or its source, was necessarily perceived as legitimate or the last word, verification was always a contest of elites vs elites. The concept of expertise itself was never challenged, and always perceived as legitimate. If, as Reich and Barnoy (2016) believe, the boundary between news creators and consumers is increasingly porous, the boundary for statistics creation remains impermeable. Putting it another way, there were no citizen-generated statistics. This is additional evidence that perceptions of interactivity in journalism are running ahead of its real effects, which Anderson (2013) and Domingo (2008) found were only limited. Both these authors concern themselves with the professional authority of journalists to set media content vs the authority of non-journalists (a jurisdictional issue) or with the question of what kind of expertise in fact-finding journalists possess, without asking about the authority or expertise behind the knowledge itself. The authority of numbers may owe something to the continued power of what Strathern (2000) calls "audit culture," the tendency to define knowledge solely in terms of what can be measured. In other words, the most relevant question may not be the tussle between citizens and professionals, but the underlying question of why the authority of statistics and their highly trained and credentialed creators is considered legitimate in the first place and never contested.

This study also fails to explain the unexpected appearance of subjects who, contrary to the general pattern, did *not* ignore the conceptual issues behind statistics construction. Recognizing that the truth of a scientific finding cannot be separated from the way it was investigated, three subjects defied widespread journalistic norms to maintain a continuing interest in scientific methodology even when it was not relevant to any particular story and occasionally used this insight to report on the quality of individual research findings. Subject 1, who had partly lost interest in the restrictions of daily journalism had followed some science closely enough to be critical of some of its methodology.

Previous research (Bell, 1994; Crow and Stevens, 2012; Dunwoody, 2004; Giannoulis, Botetzagias, and Skanavis, 2010; Gibson et al, 2016; Mcinerney, Bird, and Nucci, 2004; Vestergård, 2011; Wilson, 2000, 2002) did not find this kind of training in the toolkit of most science journalists, making the three subjects "unicorns" by most science journalism standards. Figdor (2017) believed most science journalists were not equipped to evaluate the quality of scientific knowledge and research on their own. Findings such as the above call that into question. Moreover, three subjects out of 15 (20 per cent) represents a non-negligible fraction, making these findings more difficult to dismiss. None of the three subjects found it easy to escape the familiar routines of science reporting, including a disinclination to go too deeply "into the weeds." However, subject 11 reported some success publishing stories about problems with scientific methods. At present the breadth of this kind of methodological insight among science journalists remains to be determined.

Third, it does not fully plumb the ways in which statistics only available from a single source may shape journalists' perception of them. Subjects repeatedly justified using some statistics, such as pollution levels, for lack of alternatives. This monopoly power probably derives from the

historically exclusive position of the state in creating certain kinds of quantified knowledge such as censuses and weather data (Kruger, Daston and Heidelberger, 1987; Porter, 1986). Does this monopoly over the creation of certain kinds of knowledge combine with the monopoly power of the state's political authority to give such statistics a special status in the culture, which reaches journalists through the hierarchy of influences process? That government statistics are usually considered reliable and professionalized does not insulate them from politics or from problems with their conceptual basis (Fioramonti, 2013; Merry, 2016; Porter, 1996). In fact Lugo-Ocando and Brandão (2016) and Lugo-Ocando and Lawson (2017) say government agencies tailor their statistics not only for good measurement but also with an eye toward conceptualizing the things being measured in ways that will gain attention for the work of the agencies that create them.

Regardless of whether individual journalists are aware of this doubleness, the routinized production of statistics makes them easy to integrate into the routines of journalism, giving them the status of a *de facto* information subsidy that serves the needs of both the journalists (who need the numbers) and the authorities who create them (who can use particular concepts of a number to reinforce one view of an issue over another). This suggests that when certain kinds of ideological biases in persist statistics, they persist not for reasons that are both structural and ideological, in ways that are difficult to tease apart.

Future research should address this combination of self-reinforcing social arrangements, which probably makes statistical monopolies hard to change, contributing to the perception that their products are unarguably "real." Lugo-Ocando and Lawson (2017) say statistics-creating agencies do not distinguish between what might be called their "sacred" purposes (public knowledge) and their "secular" purposes (favoring one policy over another, with its attendant need to make particular concepts of a phenomenon more visible than others). Creating

knowledge and legitimizing policy are inseparable in this view, with media needs a fundamental part of the process. The consistency of this study's findings, and their generalizability, suggests they could be the basis for studying the monopoly issue and its multiple causes as a phenomenon in itself.

Recommendations for future action

The findings in this paper have implications not only for journalism but for the audiences journalists serve. How would the public think about the facts they take for granted if journalists were freed from reductive and essentialist ideas about statistics and brought more of the professional habit of learned skepticism to all the numbers that increasingly govern not only the news, but life in general? This is an easier question to ask than to answer. The family of practices broadly labeled data journalism come to mind as an obvious place to begin, but even there the issue isn't simple. What Coddington (2015) called "journalism's quantitative turn" (p. 331) has potential to bring public data to public account and make democracy more responsive. Yet data journalism remains poorly defined and undeveloped, with multiple ways of using and integrating data into editorial product and an increased collaboration between traditionally experienced journalists and non-journalists in the newsroom such as programmers. Adding further instability, quantitative journalism exists alongside and partly overlaps the ongoing challenge posed by citizen journalism and open sourcing to traditional professional roles and ideas of professional expertise. How do data or non-data journalists learn the benefits of data literacy? Will the regular work of sourcing, handling and analyzing data eventually teach data journalists a deeper understanding of the ways seemingly neutral and objective statistics are constructed and what they mean and don't mean? What about more traditional journalists such as the subjects of this

study who sometimes recognized the conceptual basis for data creation, sometimes checked statistics through comparative methods that ignored conceptual issues, and sometimes passively accepted statistics on the say-so of authorities--an amalgam of practices that could not be "simply and smoothly reduced into rubrics and categories"? (Reich and Barnoy, 2016 p. 2)

One way to begin changing this stalemate might be to meet the profession where it already stands by developing a mixed-approach course in statistics specifically aimed at journalists. This could combine 1) the "internal" conceptual aspects of frequentist statistics such as frequency distributions and correlations, along with 2) their "external" aspects such as the ways statistics are conceptualized, and the ideas, arguments and conflicts behind these, and the material aspects of the counting process itself, which can exert their own influence over finished statistical products.

This proposal has several advantages currently missing from journalism education. First, it engages those whom it teaches by immediately connecting journalism students (and experienced journalists) with aspects of quantification they may already recognize through their past engagement with official sources of statistics such as government budgets or unemployment figures. This is particularly true for the so-called "number junkies" or "number geeks" (Van Witsen, 2018), professionals who have already been involved deeply enough with statistics-related stories to have had some contact with the actual creators and learned enough about their methods to become intrigued by them. Connecting this aspect of journalists' practice-derived knowledge to the conceptual principles behind number creation (Alonso and Starr, 1987; Andreas and Greenhill, 2010) can reinforce their understanding of each one by showing them both how principles function in actual practice and how the particularities of statistical practice are driven by powerful but latent principles. Second, by letting journalists experiment with

numbers derived from their own reporting, it allows them to grasp how data they are familiar with in practice connect to the internal aspects of numbers handling such as the ways that means, medians or standard deviations emerge from a mass of statistics in previously covered stories and change as the numbers in those stories change over time. Third, by giving students close, hands-on acquaintance with the actual processes of counting and measuring (such as a guest talk by a manager of a statistical agency or a site visit with a data collector) it can teach them how to distinguish the idealized processes of data handling and analysis taught in texts from the surprises, conundrums and contingencies familiar to those who do the actual work of abstracting the unevenness of events or experiences into uniform data. This can familiarize students with the importance of commensuration (Bowker and Star, 2000; Espeland and Stevens, 1998) the critical but rarely recognized challenge of fitting countless events that are alike to a greater or lesser degree into previously existing categories in order to conduct a uniform count. Espeland and Stevens (1998) say different ways of commensurating things can have different policy effects (such as deciding what counts as unemployment vs underemployment) and therefore function as instruments of policy and power. Learning this can be an important link between the abstract processes of counting and journalists' familiar newsgathering practice.

Teaching statistics this way, in addition to engaging journalists where they already "live," might also overcome some of the longstanding challenges faced by statistics education in journalism. Program heads frequently believe (Dunwoody and Griffin, 2013) faculty are not always equipped to teach statistics and students not always interested in learning. Dunwoody and Griffin found the incentives some program leaders offered did not change the level of statistical instruction given in journalism programs over the course of eleven years,

suggesting the problem may lie not in leadership or rewards but in something about the subject of statistics itself or the way it has been taught.

This is a fraught subject, with frequent accusations by journalists and non-journalists both, that journalists lack something called "numeracy," an imprecisely conceptualized term (Curtin and Maier, 2001; Harrison, 2016; Nguyen and Lugo-Ocando, 2016) that appears, broadly, to mean familiarity with, aptitude for, or comfort with numerical concepts and operations. Putting accusations aside, Harrison (2016) administered a math quiz to 32 journalism students and 40 statistics students. Results were mixed, with statistics students strongly outpacing journalism students on three questions out of ten, journalism students beating statisticians on one, and close matches on six more. This is neither a triumph nor a defeat for the journalists, leading the author to consider whether the stereotype of journalists' poor aptitude for numbers might be a self-perpetuating myth. Separating questions of performance from questions of ability, he theorized that journalists' performance with numbers might be due to situational factors such as time pressure, unreliable source information or errors between different levels of the news production process, or caused by poor self-image with math rather than lack of innate math talent, a view supported by Curtin and Maier (2001).

Harrison also found that the division between narrative and non-narrative knowledge was taken for granted, contributing to the myth by making it easier for students to think of themselves as word- or numbers-oriented. This appears to be borne out by this study, particularly in the way most subjects were deeply engaged with numbers, yet not according to accepted statistical methods, and the way some had an intuitive grasp of the conceptual issues behind numbers creation even if they applied that understanding inconsistently. These findings support idea that some of journalism's problems with statistics lie less in technique than in perspective; therefore

the teaching of statistics to journalists should derive from the way their existing knowledge actually functions. The course in data literacy described above links this knowledge about journalists' actual practices with statistics (and the underlying ideas behind them) to principles of both statistical operations and statistical conceptualization to help journalists move easily and comfortably between the two realms of knowledge. Curtin and Maier (2001, p. 734) say, "many journalists remain convinced that numbers cannot—and should not be—their calling. Researchers can make worthy contributions by showing how to change these deeply embedded social norms." The findings of this dissertation might be a starting point for this task.

APPENDICES

APPENDIX A: SEMISTRUCTURED INTERVIEW QUESTIONS

Potential interview questions:

GROUP ONE
Ideas about statistics

How objective are numbers to you?
Are statistics objective?
How much knowledge do you need to have to report on numbers?
What do you look for? Clarity? Relevance? Simplicity? A single statistic that contains a great deal of
meaning?

GROUP TWO
Specifics about individual stories or statistics

Are you a staff reporter or freelancer?
What beat?
Computational reporting or traditional source-based reporting?

(Referring to specific story)
How did you decide to include this particular number?
Why did you consider this number worth including?
Was the number or the source checked?
Why or why not?
What in particular led you to believe this number?
How often have you used each source in the past?
Which sources were entirely new to this story and which were repeat sources?
Did you check this number with another source?
How did you find alternative sources?
What were you hoping to gain from verifying?
Was that hope justified?

GROUP THREE
General questions about statistics in news stories

How did you decide to treat this as a hard news story and not follow up on it?
Can you think of a time you used a lot of stats?
Was it in the source material?

GROUP FOUR
Ideas about how statistics function in the news

What signals to you when to trust numbers?

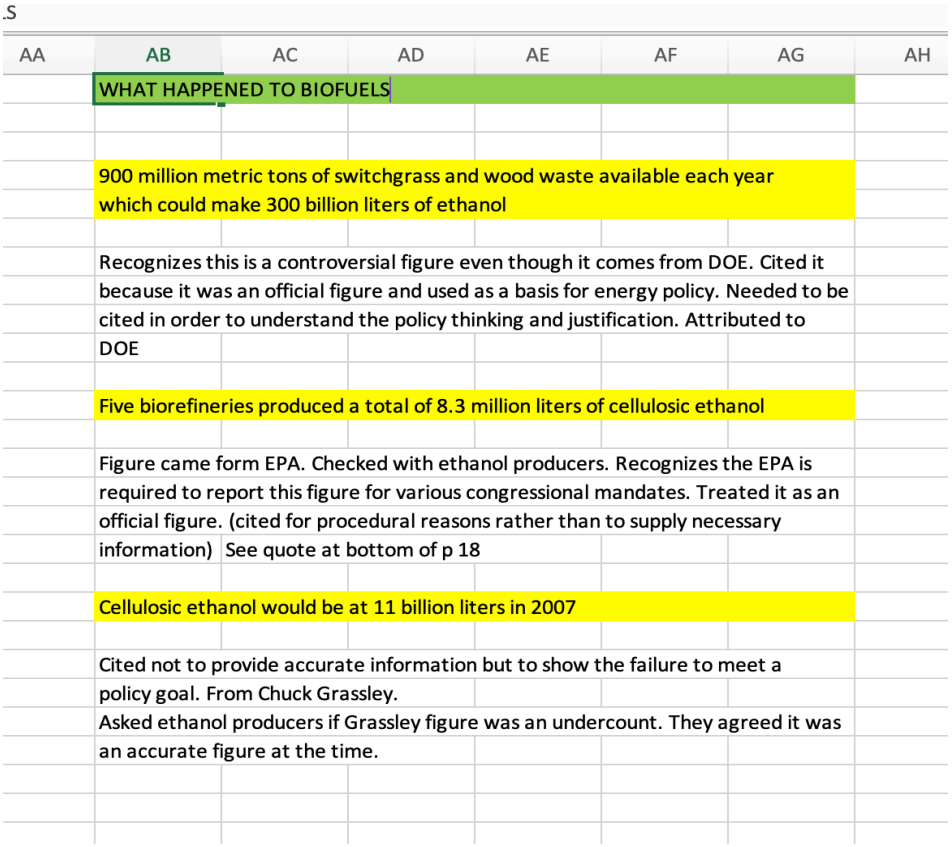
When do you inquire into where a number comes from?
Which numbers need justification
Which numbers speak for themselves
What triggers your instinct to check something out or to go with limited or single source?
What constitutes an acceptable quality crosscheck
What priorities do you bring to hard news v features?
How do you decide how many statistics to include?
Do you make different choices for health stories where there's a danger of raising false hopes, than for non health-related science stories?
How do you decide how many sources to consult?
What is the value of different kinds of sources? Including:

- Scientists/scientific institutions
- NGOs
- Lobbying or interest groups

Where do you look for other or alternate sources? Other journals? Other scientists?

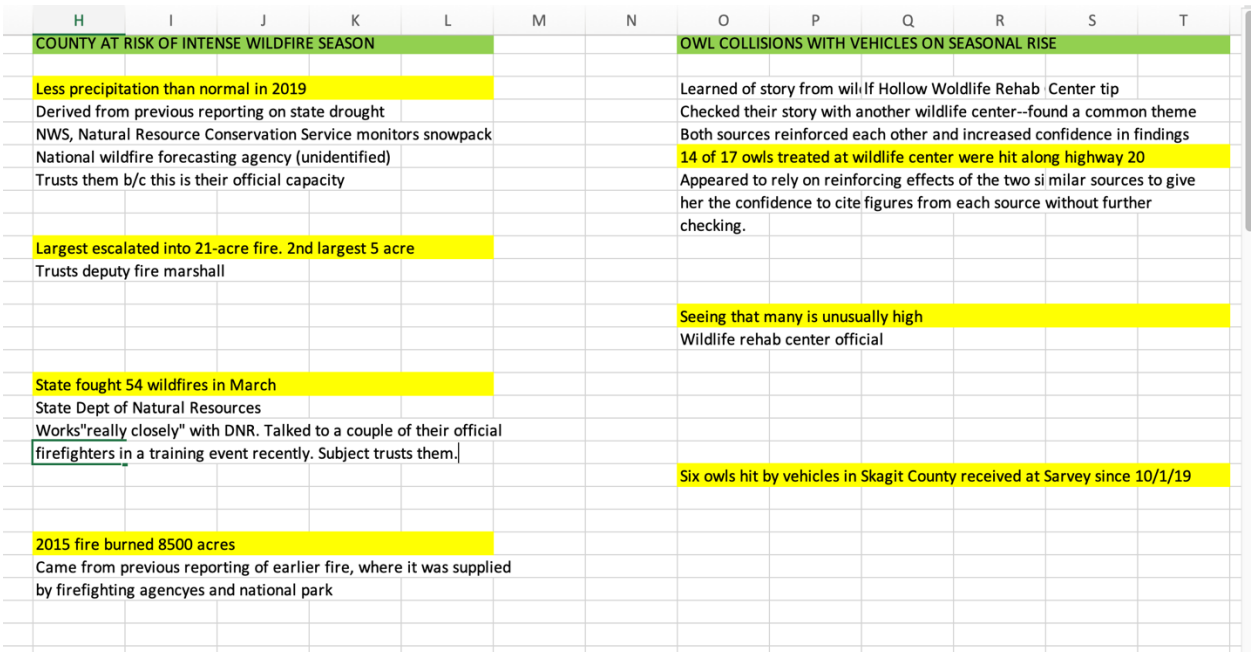
How important is it to news stories to have numbers in them? How interested do you think readers are in numbers? Or in this particular number?
Do you think differently for statistics in hard news stories on a short deadline than for statistics in longer timeframe stories?
Do you ever ask how a survey was done? When do you decide to include something about methodology?
How do you know when to trust a survey?
Is there any kind of test for when statistics might be deceptive? What triggers your suspicion?

APPENDIX B: EXCEL SCREENSHOT SAMPLES



AA	AB	AC	AD	AE	AF	AG	AH
	WHAT HAPPENED TO BIOFUELS						
	900 million metric tons of switchgrass and wood waste available each year which could make 300 billion liters of ethanol						
	Recognizes this is a controversial figure even though it comes from DOE. Cited it because it was an official figure and used as a basis for energy policy. Needed to be cited in order to understand the policy thinking and justification. Attributed to DOE						
	Five biorefineries produced a total of 8.3 million liters of cellulosic ethanol						
	Figure came form EPA. Checked with ethanol producers. Recognizes the EPA is required to report this figure for various congressional mandates. Treated it as an official figure. (cited for procedural reasons rather than to supply necessary information) See quote at bottom of p 18						
	Cellulosic ethanol would be at 11 billion liters in 2007						
	Cited not to provide accurate information but to show the failure to meet a policy goal. From Chuck Grassley.						
	Asked ethanol producers if Grassley figure was an undercount. They agreed it was an accurate figure at the time.						

Figure 1: Screenshot 1



H	I	J	K	L	M	N	O	P	Q	R	S	T
COUNTY AT RISK OF INTENSE WILDFIRE SEASON							OWL COLLISIONS WITH VEHICLES ON SEASONAL RISE					
Less precipitation than normal in 2019							Learned of story from wil	lf Hollow Woldlife Rehab		Center tip		
Derived from previous reporting on state drought							Checked their story with another wildlife center--found a common theme					
NWS, Natural Resource Conservation Service monitors snowpack							Both sources reinforced each other and increased confidence in findings					
National wildfire forecasting agency (unidentified)							14 of 17 owls treated at wildlife center were hit along highway 20					
Trusts them b/c this is their official capacity							Appeared to rely on reinforcing effects of the two si	milar sources to give				
							her the confidence to cite figures from each source without further					
							checking.					
Largest escalated into 21-acre fire. 2nd largest 5 acre												
Trusts deputy fire marshall												
							Seeing that many is unusually high					
							Wildlife rehab center official					
State fought 54 wildfires in March												
State Dept of Natural Resources												
Works"really closely" with DNR. Talked to a couple of their official												
firefighters in a training event recently. Subject trusts them.												
							Six owls hit by vehicles in Skagit County received at Sarvey since 10/1/19					
2015 fire burned 8500 acres												
Came from previous reporting of earlier fire, where it was supplied												
by firefighting agencyes and national park												

Figure 2: Screenshot 2

BIBLIOGRAPHY

BIBLIOGRAPHY

Adams, W. C. (2015). Conducting semi-structured interviews. Ch. 19 in Newcomer, K. E., Hatry, H. P., and Wholey, J. S. *Handbook of practical program evaluation*, Wiley, *492-505*.

Agren, D. (18 July 2016). 'City Mexico cuts poverty at a stroke – by changing the way it measures earnings'. *The Guardian*. https://www.theguardian.com/world/2016/jul/18/mexico-cuts-povertynational-statistics-changes-earnings-measurements

Ahmad, M. I. (2016). The magical realism of body counts: How media credulity and flawed statistics sustain a controversial policy. *Journalism, 17*(1), 18-34. doi:10.1177/1464884915593237

Allan, S. (2004). *News culture*. Maidenhead; New York: Open University Press.

Alonso, W., and Starr, P. (Eds.). (1987). *The politics of numbers*. Russell Sage Foundation.

Amberg, S. M., and Hall, T. T. (2010). Precision and rhetoric in media reporting about contamination in farmed salmon. *Science Communication, 32*(4), 489-513.

Anderson, C. W. (2013). What aggregators do: Towards a networked concept of journalistic expertise in the digital age. *Journalism*, *14*(8), 1008-1023.

Anderson, C. W. (2018). *Apostles of certainty: Data journalism and the politics of doubt*. Oxford University Press.

Andreas, P., and Greenhill, K. M. (Eds.). (2010). *Sex, drugs, and body counts: The politics of numbers in global crime and conflict*. Cornell University Press.

Angwin, J., Larson, J., Mattu, S., and Kirchner, L. (2016). Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*, 23 May 2016.

Barnoy, A., and Reich, Z. (2019). The when, why, how and so-what of verifications. *Journalism Studies*, *20*(16), 2312-2330.

Bell, A. (1994). Media (mis) communication on the science of climate change. *Public understanding of science, 3, 259-275*.

Berman, E. P., and Milanes-Reyes, L. M. (2013). The Politicization of Knowledge Claims: The "Laffer Curve" in the U.S. Congress. *Qualitative Sociology, 36*(1), 53-79. doi:10.1007/s11133-012-9242-4

Best, J. (1987). Rhetoric in Claims-Making: Constructing the Missing Children Problem. *Social Problems, 34*(2), 101-121. doi:10.2307/800710

Bhatti, Y., and Pedersen, R. T. (2015). News reporting of opinion polls: Journalism and statistical noise. *International Journal of Public Opinion Research*, *28*(1), 129-141.

Bodemer, N., Meder, B., and Gigerenzer, G. (2014). Communicating relative risk changes with baseline risk: Presentation format and numeracy matter. *Medical Decision Making*, *34*(5), 615-626.

Bødker, H., and Neverla, I. (2012). Introduction: environmental journalism. *Journalism Studies*, *13*(2), 152-156.

Boellstorff, T. (2013). Making big data, in theory. *First Monday, 18*(10). doi:10.5210/fm.v18i10.4869

Bowker, G. C., and Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.

Boyd, D., and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication and Society*, *15*(5), 662-679.

Brand, R. (2008). The numbers game: A case study of mathematical literacy at a South African newspaper. *Communicatio, 34*(2), 210-221. doi:10.1080/02500160802456155

Brandão, R. (2016). *Study Shows: How Statistics Are Used to Articulate and Shape Discourses of Science in the Newsroom* (Doctoral dissertation, University of Sheffield). http://etheses.whiterose.ac.uk/16438/1/Brandão%2C%20Renata_PhD.pdf

Brandão, R., and Nguyen, A. (2017). Statistics in science journalism: An exploratory study of four leading British and Brazilian newspapers. Ch. 5 in Nguyen, A. (Ed.). *News, Numbers and Public Opinion in a Data-Driven World*. Bloomsbury Publishing. 78-92.

Brinkmann, S. (2014). Unstructured and semi-structured interviewing. Ch. 14 in Leavy, P. (Ed.). *The Oxford Handbook Of Qualitative Research.* Oxford University Press, 277-299.

Brod, M., Tesler, L. E., and Christensen, T. L. (2009). Qualitative research and content validity: developing best practices based on science and experience. *Quality Of Life Research*, *18*(9), 1263.

Brüggemann, M. (2013). Transnational trigger constellations: Reconstructing the story behind the story. *Journalism*, *14*(3), 401-418.

Callison, C., Gibson, R., and Zillmann, D. (2009). How to report quantitative information in news stories. *Newspaper Research Journal*, *30*(2), 43-55.

Carey, J. (Ed.). (1988). *Media, myths, and narratives: Television and the press.* Sage.

Casper, M. J., and Clarke, A. E. (1998). Making the Pap smear into the 'Right Tool' for the job: cervical cancer screening in the USA, circa 1940-95. *Social Studies Of Science*, *28*(2), 255-290.

Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal Of The Learning Sciences*, *6*(3), 271-315.

Choi, J., Hecht, G. W., and Tayler, W. B. (2012). Lost in translation: The effects of incentive compensation on strategy surrogation. *The Accounting Review*, *87*(4), 1135-1163.

Coddington, M. (2015). Clarifying journalism's quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, *3*(3), 331-348.

Cohn, V., and Cope, L. (2012). *News and numbers: A writer's guide to statistics*. 3rd ed. Malden MA: Wiley-Blackwell.

Conk, M.A. (1989). The 1980 census in historical perspective. Ch. 4 in Alonso, W., and Starr, P. (Eds.). *The Politics Of Numbers*. Russell Sage Foundation. 155-186.

Cottle, S. (2000). New(s) times: Towards a 'second wave" of news ethnography. *Communication,* (25), 19-41.

Crow, D. A., and Stevens, J. R. (2012). Local science reporting relies on generalists, not specialists. *Newspaper Research Journal*, 33(3), 35–48.

Curtin, P. A., and Maier, S. R. (2001). Numbers in the newsroom: A qualitative examination of a quantitative challenge. *Journalism and Mass Communication Quarterly, 78*(4), 720-738. doi:10.1177/107769900107800407

Della Porta, D., and Diani, M. (2006). *Social movements: An introduction.* Malden, MA: Blackwell.

DeSantos, M. (2009). Fact-totems and the statistical imagination: The public life of a statistic in Argentina 2001. *Sociological Theory, 27*(4), 466-489. doi:10.1111/j.1467-9558.2009.01359.x

Deuze, M. (2008). The changing context of news work: Liquid journalism for a monitorial citizenry. *International Journal of Communication*, 2(18), 848-865.

Dickinson, R. (2008) Studying the sociology of journalists: The journalistic field and the news world. *Sociology Compass* 2/5: 1383–1399, 10.1111/j.1751-9020.2008.00144.x

Diekerhof, E., and Bakker, P. (2012). To check or not to check: An exploratory study on source checking by Dutch journalists. *Journal of Applied Journalism and Media Studies*, *1*(2), 241-253.

Domingo, D. (2008). Interactivity in the daily routines of online newsrooms: Dealing with an uncomfortable myth. *Journal Of Computer-Mediated Communication*, *13*(3), 680-704.

Donsbach, W. (2014). Journalism as the new knowledge profession and consequences for journalism education. *Journalism Studies*, 15(6), 661–677.

Dunwoody, S. (2004). How valuable is formal science training to science journalists? *Comunicação e Sociedade, 6*, 75-87. doi:10.17231/comsoc.6(2004).1229

Dunwoody, S., and Griffin, R. J. (2013). Statistical reasoning in journalism education. *Science Communication*, *35*(4), 528-538.

Ekström, M. (2002) Epistemologies of TV journalism: A theoretical framework. *Journalism*, 3(3): 259–282.

Ericson, R. V. (1998). How journalists visualize fact. *The Annals of the American Academy of Political and Social Science*, *560*(1), 83-95.

Ericson, R.V., Baranek, P.M., and Chan, J.B.L. (1987). *Visualizing deviance: A study of news organizations*. Toronto, ON, Canada: University of Toronto Press.

Espeland, W. N., and Stevens, M. L. (1998). Commensuration as a social process. *Annual Review Of Sociology*, *24*(1), 313-343.

Espeland, W. N., and Stevens, M. L. (2008). A sociology of quantification. *European Journal of Sociology/Archives Européennes de Sociologie*, *49*(3), 401-436.

Ettema, J.S., and Glasser, T.L. (1985). On the epistemology of investigative journalism. *Communication, 8*(2): 183–206.

Ettema, J.S., and Glasser, T.L. (1998). *Custodians of conscience: Investigative journalism and public virtue*. New York: Columbia University Press.

Fahnestock, J. (1986). Accommodating science: The rhetorical life of scientific facts. *Written Communication*, *3*(3), 275-296.

Figdor, C. (2017). (When) is science reporting ethical? The case for recognizing shared epistemic responsibility in science journalism. *Frontiers in Communication*, *2*(3), 1-7. doi: 10.3389/fcomm.2017.00003

Fioramonti, L. (2013). *Gross domestic problem: The politics behind the world's most powerful number*. London: Zed Books.

Fishman, M. (1980). *Manufacturing the news*. Austin: University of Texas Press.

Foley, K. E. (2020). MODEL BEHAVIOR; The IHME model was built to do one job. It's hard to repurpose it for another. *Quartz,* April 22, 2020. https://qz.com/1840186/what-the-ihme-covid-19-model-can-and-cant-tell-the-us/

Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., and Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health*, *25*(10), 1229-1245.

Fusch, P.R. and Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *Walden University Scholar Works. 20*(9), 1408-1416.

Gans, H. (2004). *Deciding what's news: A study of CBS evening news, NBC nightly news, Newsweek, and Time*. Evanston, IL: Northwestern University Press

Giannoulis, C., Botetzagias, I., and Skanavis, C. (2010). Newspaper reporters' priorities and beliefs about environmental journalism: An application of Q-methodology. *Science Communication*, *32*(4), 425–466.

Gibson, T. A., Craig, R. T., Harper, A. C., and Alpert, J. M. (2016). Covering global warming in dubious times: Environmental reporters in the new media ecosystem. *Journalism*, *17*(4), 417-434.

Giddens, A, (1986). *The Constitution of society: Outline of the theory of structuration*. Berkeley: University of California Press.

Gitelman, L. (Ed.). (2013). *Raw data is an oxymoron*. MIT press.

Godler, Y., and Reich, Z. (2017). Journalistic evidence: Cross-verification as a constituent of mediated knowledge. *Journalism*, *18*(5), 558-574.

Godler, Y., and Reich, Z. (2017b). News cultures or "epistemic cultures"? Theoretical considerations and empirical data from 62 countries. *Journalism Studies*, *18*(5), 666-681.

Goldman, A. (2002). What is social epistemology? A smorgasbord of projects. In: Goldman, A. (Ed.). *Pathways to knowledge: private and public*. New York: Oxford University Press, 182–204.

Goldman, A. (2010). Why social epistemology is real epistemology. In: Haddock, A., Millar, A., and Pritchard, D. (Eds.) *Social Epistemology*. New York: Oxford University Press, pp. 1–29.

Guba, E. G., and Lincoln, Y. S. (1994). Competing paradigms in qualitative research. Ch. 6 in Denzin, N.K., and Lincoln, Y. S. (Eds.) *Handbook Of Qualitative Research*, Thousand Oaks: Sage. 105-117.

Guest, G., Bunce, A., and Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59-82.

Hacking, I. (2006). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press.

Hacking, I., and Hacking, T. (1990). *The taming of chance* (Vol. 17). Cambridge University Press.

Hand, D. J. (2009). Modern statistics: the myth and the magic. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(2), 287-306.

Harding, S. (1986). *The science question in feminism*. Cornell University Press.

Harding, S. (2016). *Whose science? Whose knowledge?: Thinking from women's lives*. Cornell University Press.

Harrison, S. (2016). Journalists, numeracy and cultural capital. *Numeracy, 9*(2). doi:10.5038/1936-4660.9.2.3

Hesse-Biber, S. N., and Leavy, P. (2010). *The practice of qualitative research*. Sage.

Howe, C. Z. (1988). Using qualitative structured interviews in leisure research: Illustrations from one case study. *Journal of Leisure Research*, *20*(4), 305-323.

Jencks, C. (1989). The Politics of Income Measurement. Ch. 2 in Alonso, W., and Starr, P. (Eds.) *The Politics Of Numbers*. Russell Sage Foundation. 83-132.

Koetsenruijter, A. W. M. (2009). How numbers make news reliable. Ch. 10 in Dam, L., Holmgreen, L. L., and Strunck, J. (Eds.). *Rhetorical Aspects Of Discourses In Present-Day Society*. Cambridge Scholars Publishing. 193-205.

Koetsenruijter, A. W. M. (2011). Using numbers in news increases story credibility. *Newspaper Research Journal, 32*(2), 74-82.

Kruger, L. E., Daston, L. J., and Heidelberger, M. E. (1987). *The probabilistic revolution, Vol. 1: Ideas in history; Vol. 2: Ideas in the sciences*. The MIT Press.

Kuhn, T. S. (1963). The essential tension: Tradition and innovation in scientific research. *Scientific Creativity: Its Recognition And Development. New York: Wiley*, 341-354.

Kuhn, T. S. (1987). *Black-body theory and the quantum discontinuity, 1894-1912*. University of Chicago Press.

Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago Press.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B., and Woolgar, S. (2013). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Lewis, S. C., and Westlund, O. (2015). Big data and journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*, *3*(3), 447-466.

Lindlof, T. R., and Taylor, B. C. (2017). *Qualitative communication research methods*. Sage publications.

Lugo-Ocando, J. (2017). *Crime statistics in the news: journalism, numbers and social deviation*. Springer.

Lugo-Ocando, J., and Faria Brandão, R. (2016). Stabbing news: Articulating crime statistics in the newsroom. *Journalism Practice*, *10*(6), 715-729.

Lugo-Ocando, J., and B. Lawson. (2017). Poor numbers, poor news: The ideology of poverty statistics in the media. Ch. 4 in Nguyen, A. (Ed.). *News, Numbers and Public Opinion in a Data-Driven World*. London: Bloomsbury. 62-77.

Maier, S. R. (2002). Numbers in the news: A mathematics audit of a daily newspaper. *Journalism Studies, 3*(4), 507-519. doi:10.1080/1461670022000019191

Maier, S. R. (2003). Numeracy in the newsroom: A case study of mathematical competence and confidence. *Journalism & Mass Communication Quarterly*, *80*(4), 921-936.

Marshall, B., Cardon, P., Poddar, A., and Fontenot, R. (2013). Does sample size matter in qualitative research?: A review of qualitative interviews in IS research. *Journal of Computer Information Systems*, *54*(1), 11-22.

McCarthy, A. (2020). More thoughts on computing the COVID-19 fatality rate. *National Review,* March 27, 2020. https://www.nationalreview.com/2020/03/coronavirus-fatality-rate-computing-difficult/#slide-1

McConnell, E. D. (2014). Good news and bad news numeracy and framing in newspaper reporting of unauthorized migration estimates in Arizona. *Aztlán: A Journal of Chicano Studies, 39*(1), 41-70.

McConway, K. (2016). Statistics and the media: A statistician's view. *Journalism, 17*(1), 49-65. doi:10.1177/1464884915593243

McInerney, C., Bird, N., and Nucci, M. (2004). The flow of scientific knowledge from lab to the lay public: the case of genetically modified food. *Science Communication*, *26*(1), 44-74.

Merriam, S. (1995). What can you tell from an N of 1?: Issues of validity and reliability in qualitative research. *PAACE Journal of lifelong learning*, *4*, 51-60.

Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. University of Chicago Press.

Meyer, P. (2002). *Precision journalism: A reporter's introduction to social science methods*. Rowman and Littlefield.

Moynihan, R., Bero, L., Ross-Degnan, D., Henry, D., Lee, K., Watkins, J., . . . Soumerai, S. B. (2000). Coverage by the news media of the benefits and risks of medications. *New England Journal of Medicine, 342*(22), 1645-1650. doi:10.1056/nejm200006013422206

Nguyen, A., and Lugo-Ocando, J. (2016). The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism, 17*(1), 3-17.

Nisbet, M. C., and Fahy, D. (2015). The need for knowledge-based journalism in politicized science debates. *The ANNALS of the American Academy of Political and Social Science*, 658, 223–234.

Peters, E., Vastfjall, D., Slovic, P., Mertz, C. K., Mazzocco, K., and Dickert, S. (2006). Numeracy and intuitive number sense in decision making. *Psychological Science, 17*(5), 407-413.

Petersen, W. (1989) Politics and the measurement of ethnicity. Ch. 5 in Alonso, W., and Starr, P. (Eds.) (1989). *The Politics Of Numbers*. Russell Sage Foundation. 187-234.

Porter, T. M. (1986). *The rise of statistical thinking, 1820-1900*. Princeton University Press.

Porter, T. M. (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Prewitt, K. (1987). Public statistics and democratic politics. In *The Politics Of Numbers.* Russell Sage Foundation. 261-274.

Prewitt, K. (2013). *What is your race?:The census and our flawed efforts to classify Americans*. Princeton University Press.

Reese, S. D. (2016). Theories of journalism. In *Oxford Research Encyclopedia of Communication*.

Reese, S. D., and Shoemaker, P. J. (2016). A media sociology for the networked public sphere: The hierarchy of influences model. *Mass Communication and Society*, *19*(4), 389-410.

Reich, Z. (2014). 'Stubbornly unchanged': A longitudinal study of news practices in the Israeli press. *European Journal of Communication*, *29*(3), 351-370.

Reich, Z., and Barnoy, A. (2016). Reconstructing production practices through interviewing. Ch. 32 in Witschge, T., Anderson, C. W., Domingo, D., and Hermida, A. (Eds.) *The SAGE Handbook Of Digital Journalism*. Sage. 477-493.

Reich, Z., and Barnoy, A. (2019). Disagreements as a form of knowledge: How journalists address day-to-day conflicts between sources. *Journalism*, 1-19.

Richardson, R., and Kramer, E. H. (2006). Abduction as the type of inference that characterizes the development of a grounded theory. *Qualitative Research*, *6*(4), 497-513.

Roeh, I. (1989). Journalism as storytelling, coverage as narrative. *The American Behavioral Scientist, 33*(2), 162-168.

Roeh, I., and Feldman, S. (1984). The rhetoric of numbers in front-page journalism: how numbers contribute to the melodramatic in the popular press. *Text - Interdisciplinary Journal for the Study of Discourse*, *4*(4), 347-368.

Rose, N. (1991). Governing by numbers: Figuring out democracy. *Accounting, Organizations and Society, 16*(7), 673-692. doi:10.1016/0361-3682(91)90019-b

Scheufele, D. A., Krause, N. M., Freiling, I, and Brossard, D. (2020). How not to lose the COVID-19 communication war. *Issues in Science and Technology* (April 17, 2020).

Schudson, M. (1982). The politics of narrative form: the emergence of news conventions in print and television. *Daedalus, 111*(4), 97-112.

Schudson, M. (1989). The sociology of news production. *Media, Culture & Society* 11, 263-282.

Schultz, I. (2007). The journalistic gut feeling: Journalistic doxa, news habitus and orthodox news values. *Journalism Practice, 1*(2), 190-207.

Sigal L. (1973.) *Reporters and officials: the organization and politics of newsmaking*. Lexington, MA: Lexington Books

Simmerling, A., and Janich, N. (2016). Rhetorical functions of a 'language of uncertainty' in the mass media. *Public Understanding of Science*, *25*(8), 961-975.

Steinhardt, S. (2019). Technoscience in the era of #MeToo and the science march. *Journal of Science Communication*, *18*(4), 1-5.

Stonbely, S. (2015). The social and intellectual contexts of the U.S. "Newsroom Studies," and the media sociology of today. *Journalism Studies, 16*(2), 259-274.

Strathern, M. (2000). *Audit cultures: anthropological studies in accountability, ethics, and the academy*. London: Routledge.

Tooze, J. A. (2001). *Statistics and the German state, 1900-1945:The making of modern economic knowledge* (Vol. 9). Cambridge University Press.

Tuchman, G. (1972). Objectivity as strategic ritual: an examination of newsmen's notions of objectivity. *American Journal of Sociology, 77(*4), 660-679.

Tuchman, G. (1973). Making news by doing work: Routinizing the unexpected. *American Journal of Sociology*, *79*(1), 110-131.

Tufekci, Z. (2020). Don't believe the COVID-19 models. *Atlantic,* April 2, 2020. https://www.theatlantic.com/technology/archive/2020/04/coronavirus-models-arent-supposed-be-right/609271/

Tunstall, J. (1972). News organization goals and specialist newsgathering journalists. Ch. 12 in McQuail, D. (Ed.). *Sociology of Mass Communications: Selected Readings.* Harmondsworth: Penguin Books Ltd. 259-280.

Usher, N. (2013). Marketplace public radio and news routines reconsidered: Between structures and agents. *Journalism*, *14*(6), 807-822.

van Rijnsoever, F. J. (2017). (I can't get no) saturation: a simulation and guidelines for sample sizes in qualitative research. *PLoS One*, *12*(7).

Van Witsen, A. (2018). How journalists establish trust in numbers and statistics: Results from an exploratory study. In K.P. Hunt (Ed.). *Understanding the Role of Trust and Credibility in Science Communication: Proceedings from the 6th ISU Symposium on Science Communication.* .https://doi.org/10.31274/sciencecommunication-181114-8

Van Witsen, A. (2019). How daily journalists use numbers and statistics: the case of global average temperature. *Journalism Practice.* 1-19. https://doi.org/10.1080/17512786.2019.1682944

Vestergård, G. L. (2011). From journal to headline: The accuracy of climate science news in Danish high quality newspapers. *JCOM: Journal of Science Communication*, *10(*2), 1-7.

Volante, L. (2004). Teaching to the test: what every educator and policy-maker should know. *Canadian Journal of Educational Administration and Policy,* 35, 1-6. https://files.eric.ed.gov/fulltext/EJ848235.pdf

Warren, K. B. (2010). The illusiveness of counting "victims" and the concreteness of ranking countries: trafficking in persons from Colombia to Japan. Ch. 5 in Andreas, P., and Greenhill, K. M. (Eds.). *Sex, drugs, and body counts: The politics of numbers in global crime and conflict*. Cornell University Press. 110-126

Weitzer, R. (2007). The social construction of sex trafficking: Ideology and institutionalization of a moral crusade. *Politics & Society*, *35*(3), 447-475.

Wilson, K. M. (2000). Drought, debate and uncertainty: measuring reporters' knowledge and ignorance about climate change. *Public Understanding of Science,* 9, 1-13.

Wilson, K. M. (2002). Forecasting the future: How television weathercasters' attitudes and beliefs about climate change affect their cognitive knowledge on the science. *Science Communication, 24*(2), 246-268.

Yong, E. (2020). Why the coronavirus is so confusing. *Atlantic,* April 29, 2020. https://www.theatlantic.com/health/archive/2020/04/pandemic-confusing-uncertainty/610819/